

ROBIN MANSELL, FLAVIA DURACH, MATTHIAS KETTEMANN,
THÉOPHILE LENOIR, ROB PROCTER, GYAN PRAKASH TRIPATHI, & EMILY TUCKER

INFORMATION ECOSYSTEMS AND TROUBLED DEMOCRACY

THE STATE OF KNOWLEDGE ON NEWS
MEDIA, AI, AND DATA GOVERNANCE

NORDICOM



OBSERVATORY ON
INFORMATION AND
DEMOCRACY



Nordicom is funded with
support from the
Nordic Council of Ministers

This book was produced with support from the Observatory on Information and Democracy,
and published in collaboration with Nordicom at the University of Gothenburg.

Information Ecosystems and Troubled Democracy

The State of Knowledge on News Media,
AI, and Data Governance

Information Ecosystems and Troubled Democracy: The State of Knowledge on News Media, AI, and Data Governance

Robin Mansell, Flavia Durach, Matthias Kettemann,
Théophile Lenoir, Rob Procter, Gyan Prakash Tripathi,
& Emily Tucker

© 2025 Respective authors. This work is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public licence (CC BY-NC-ND 4.0). To view a copy of the licence, visit creativecommons.org/licenses/by-nc-nd/4.0/

ISBN (PRINT): 978-91-89864-22-1

ISBN (PDF): 978-91-89864-23-8

ISBN (EPUB): 978-91-89864-24-5

This work is Open Access and available to read free of charge and without registration requirement at www.nordicom.gu.se/en/publications/information-ecosystems-and-troubled-democracy

DOI: 10.48335/9789189864238

CITE THIS WORK:

Mansell, R., Durach, F., Kettemann, M., Lenoir, T., Procter, R., Tripathi, G. P., & Tucker, E. (2025). *Information ecosystems and troubled democracy: The state of knowledge on news media, AI, and data governance*. Nordicom, University of Gothenburg. <https://doi.org/10.48335/9789189864238>

PUBLISHED BY:

Nordicom, University of Gothenburg
Box 713
SE 405 30 Göteborg, Sweden

www.nordicom.gu.se/en

NORDICOM EDITORIAL TEAM

Manuscript editor: Kristin Clay
Managing editor: Josefine Bové
Cover by: Karin Andén

GRAPHIC DESIGN BY: Henny Östlund

FIGURES DESIGNED BY: Matthieu Mercier, adapted by Nordicom

INFORMATION ECOSYSTEMS AND TROUBLED DEMOCRACY

THE STATE OF KNOWLEDGE ON NEWS MEDIA, AI, AND DATA GOVERNANCE

ROBIN MANSELL, FLAVIA DURACH, MATTHIAS KETTEMANN,
THÉOPHILE LENOIR, ROB PROCTER, GYAN PRAKASH TRIPATHI,
& EMILY TUCKER

In memory of

Christophe Deloire

late Chair of the International Forum on Information and Democracy

Contents

Preface..... 9

Foreword..... 10

List of tables and figures..... 12

Chapter 1. Information ecosystems and democracy 13

Introduction 13

Information ecosystems and the public sphere 14

Platformisation, datafication, and democracy 17

Information integrity and mis- and disinformation 19

 Defining mis- and disinformation 20

 Measuring mis- and disinformation 22

 Human rights and governing information ecosystems 23

Scope and limitations..... 26

Conclusion 27

Chapter 2. News media, information integrity, and the public sphere..... 29

Introduction 29

Defining “news” 29

News media, the public sphere, and democracy 30

News media industry structure and platform dependency 32

Trust in news media..... 38

 Evidence of declining individual trust in news 39

 Models and concepts of media effects..... 40

 Relational models of media engagement 43

 Changing journalism standards and news media practices 45

 News media use, news avoidance, and resilience..... 47

 News media avoidance 48

 The weaponisation of online information 49

Mis- and disinformation, political polarisation, and public opinion 54

Strengthening trust and resilience to mis- and disinformation 59

Conclusion 62

Chapter 3. Artificial intelligence, information ecosystems, and democracy..... 63

Introduction 63

Defining AI..... 64

AI systems and human rights 67

 Algorithmic bias and fairness..... 68

 Freedom of expression and information..... 71

Privacy protection	71
Democracy and participatory rights.....	73
AI systems and content governance.....	74
AI systems in content generation	76
AI systems in content moderation and curation	76
AI systems and news media	80
Use of generative AI by mis- and disinformation actors	84
AI systems and democracy	87
Conclusion	92
Chapter 4. Big Tech power and data governance.....	93
Introduction	93
Information ecosystem contestations.....	93
Corporate data monopolisation and digital infrastructures	97
Data monopolisation and data dependency.....	98
Business models, knowledge, and mis- and disinformation.....	101
Big Tech monopolisation of data and AI system biases	105
Monopolising knowledge by foreclosing access to information	108
Research independence and verifiable knowledge	110
Data governance, data justice, and democratic deliberation	111
Corporate power and lobbying.....	112
Towards data justice.....	114
Conclusion	117
Chapter 5. Governing information ecosystems.....	119
Introduction	119
Varieties of information ecosystems governance.....	119
Governing principles and legislation	121
Governing network infrastructure.....	123
Governing digital platforms.....	126
Governing news media	129
Governing AI systems.....	133
Privacy and data governance	139
Governing mis- and disinformation.....	144
Safely navigating information ecosystems	149
Media and information literacy	149
AI literacy	154
Conclusion	157
Chapter 6. Governing information ecosystems in practice	159
Introduction	159
Awareness of mis- and disinformation	159

Public awareness of risks and harms.....	160
Understanding automated content moderation.....	163
Policymaker understanding of AI systems and platforms.....	165
Measures to combat mis- and disinformation	170
Public information and fact-checking.....	171
Government action: Legislation and codes of practice.....	176
Platform action to address mis- and disinformation.....	179
Alternative data governance practices	179
Epistemic privileging and data justice	182
Resistance to corporate datafication strategies.....	184
Individual digital self-defence strategies	187
Public interest alternative news media	188
Community collaborative strategies.....	189
Social entrepreneurs and community-controlled technologies.....	191
National-level decentralised data governance frameworks	194
Civil society organisations, researchers, and philanthropic organisations.....	195
Conclusion	197

Chapter 7. Conclusion: Information ecosystems and troubled

democracy.....	199
Introduction	199
Epistemic difference and inequities	200
Conceptual framings, research designs, and methodologies.....	201
Research design and methods	202
Future research priorities	204
Datafication-for-profit business models	204
Governing information ecosystems	205
The infrastructure layer	206
News media and trust.....	206
AI systems and content governance	207
Literacy and informed publics	208
Governing democratically: Towards data justice	209
Conclusion	210
Acknowledgements	211
Authors	213
Bibliography.....	214
Appendix	294

Preface

I began working with the International Observatory on Information and Democracy as scientific director for its first cycle of work in October 2023. Our mandate was to critically evaluate state-of-the-art research on information ecosystems and why changes in these ecosystems are so challenging for policymakers and many others in the Global North and the Global Majority World. My co-authors and I produced an extensive report, *Information Ecosystems and Troubled Democracy*, in January 2025.

When our project started, public concern about deepfake deceptions and the weaponisation of online information was growing. The United Nations was calling for responses to “information disorder”. Our research assessment addressed how principles for governing information ecosystems are translated into practice and the uncertainties and pressures public institutions and other actors face when they try to combat mis- and disinformation. What can be done to foster online information and communication spaces that respect human rights, do no harm, and underpin justice for all?

We were invited to present our findings to academics, civil society organisations, and policymakers in countries around the world. These meetings sparked debate about what must be done better to respond to mounting evidence that advances in digital technologies – and artificial intelligence systems especially – are boosting market leadership for some by using intrusive and exploitative data practices, while failing to safeguard both adults’ and children’s rights.

Participants in these events often told us they were incorporating our report into their teaching. When Hannu Nieminen, professor emeritus at the University of Helsinki, asked me if we were going to make the report available as a book to reach more students and policymakers, I agreed to update and revise the report. I thank Professor Nieminen for his suggestion, my co-authors, and especially professor Rob Procter, who helped with editing the revised book manuscript. Throughout our project, we all enjoyed enormous support from the Observatory and hundreds of experts.

Tackling risks and harms associated with today’s information ecosystems is urgent. This must be done in a way that suits diverse contexts. Evidence is needed to gain momentum for supporting information ecosystem governance and a public sphere that serves individuals well and supports the common good. We are enticed by Big Tech companies to depend upon their products and services and partake in the benefits of their use, but these benefits pale in comparison to the risks and harms for the excluded and disadvantaged. I hope this book provides a basis for building that momentum.

Robin Mansell
Professor Emerita
London School of Economics and Political Science
London, September 2025

Foreword

From the Arab Spring to large-scale information manipulation and disorder, Big Tech has moved from being a promising asset for a more democratic and free future to a factor in its potential doom. The harms – not only to democratic societies and processes, but also to individuals – are clear and well documented, as well as the major challenges humanity must face, including climate change, gender (in)equality, increasing migration, and peace.

During this period, civil society, journalists, and research institutions have launched a wide array of projects providing research and accountability mechanisms, seeking to have a better understanding of power dynamics in the information space and to hold the most powerful to account.

From the 2018 Cambridge Analytica scandal to now, new organisations have been created, older ones have expanded their work, and new research labs have been established to develop a deeper understanding of the scale of the issues. Hundreds of articles, reports, and books from academics, whistle-blowers, and activists have been published, providing us with important insights on information ecosystems.

Yet, as Europe was designing the first regulatory framework to respond to some of these challenges and winning the “race of the turtles”, as Maria Ressa would put it, one piece was missing: a global study to synthesise the state of knowledge on information ecosystems and their impact on societies.

Founded in November 2019, the Forum on Information and Democracy facilitates the implementation of high-level democratic safeguards for information ecosystems, enshrined in an Intergovernmental Partnership led by France. With its unique mandate from the 55 countries which have endorsed this partnership, the organisation was well positioned to do the work of critical research assessment and to build an interface between knowledge producers and policymakers.

Building on its active international community of researchers and practitioners, the Forum launched the Observatory on Information and Democracy to be equivalent to the Intergovernmental Panel on Climate Change, but for the information space. With our ambition set high, we launched in 2021 with our Prefiguration Group co-chaired by the dynamic duo of Ángel Gurría, former secretary general of the OECD, and Shoshana Zuboff, professor emerita at Harvard Business School.

Fast forward to January 2025. In the space of only a few weeks, we celebrated the publication of the first report on *Information Ecosystems and Troubled Democracy*, and we had to manage the techno-political shift in the United States and its ramifications around the world. As the information space became a hot topic, debated on television and in so many op-eds, the foundational work of our Observatory soon became a critical source of knowledge to help us go beyond the mounting news cycle attacks on information integrity. I welcome its wider accessibility in its new revised and updated book form.

Of course, neither the report nor this book, as good as they may be, can solve sociopolitical crises. It will take more research, more data, committed policymakers, and a global movement to push those who can act to defend democracy and the information and communication space. The Forum on Information and Democracy, with its projects and community, will continue to play its part.

Camille Grenier
Executive Director
Forum on Information and Democracy
Paris, September 2025

List of tables and figures

TABLE 1.1	Distinguishing between misinformation and disinformation	23
TABLE 2.1	Tensions between digital platforms and news media organisations	37
TABLE 2.2	Platform roles in the weaponisation of information (selected countries)	55
TABLE 3.1	AI system use in news media gatekeeping	84
TABLE 5.1	United Nations principles for information integrity	124
TABLE 5.2	UNESCO principles for governing digital platforms.....	128
TABLE 6.1	Understanding AI	165
TABLE 6.2	Selected mis- and disinformation countermeasures.....	173
TABLE 6.3	Corporate datafication resistance strategies and tactics.....	188
TABLE A1	Corpus source selection criteria	295
TABLE A2	Definitions of source quality.....	296
TABLE A3	Research assessment contributors	300
TABLE A4	Research Assessment Panel (RAP) contributors.....	302
FIGURE 1.1	Information ecosystems and the public sphere	17
FIGURE 2.1	Building trust in news.....	63
FIGURE 3.1	Example of realistic AI-generated face	87
FIGURE 3.2	Deepfake image of Donald Trump.....	88
FIGURE 4.1	Map of mis- and disinformation campaigns in West Africa, and state sponsorship	105
FIGURE 5.1	Simplified view of the governance of information ecosystems ...	123
FIGURE 5.2	World Press Freedom Index, 2024 (per cent)	134
FIGURE 5.3	Non-exhaustive MIL competencies	153
FIGURE 6.1	Respondents indicating they have ever used ChatGPT, by age (per cent).....	166
FIGURE 6.2	Network of fact-checking organisations, 2014–2021	175
FIGURE A1	Number of publication types, by regional focus.....	297
FIGURE A2	Number of books and book chapters, by most frequently occurring publishing company	298
FIGURE A3	Most frequently cited journals	299

Information ecosystems and democracy

“The right to know is the right to live”
(Aruna Roy, 2004: para. 1)

Introduction

In 2024, the level of democracy enjoyed by the average person globally had declined to what it was in 1985, meaning that around 72 per cent of the world’s population is living in autocracies (V-Dem Institute, 2025). On a different measure, Internet freedom declined for the 14th year in a row globally in 2024 (Freedom House, 2024). The World Economic Forum’s *Global Risks Report* ranks mis- and disinformation as the greatest risks in the coming two years (WEF, 2024, 2025) and there is “good” press freedom in only seven countries (RSF, 2025). People report rising concerns about “fake” news, including 72 per cent in the United States and 81 per cent in South Africa. In northern Europe, despite overall lower levels of concern, nearly half of the population express concern, for example, 45 per cent in Norway (Newman et al., 2024). Trust is experiencing an overall – although uneven – decline across countries, ranging from a low of 23 per cent in Greece and Hungary and 32 per cent in the United States to a high of 69 per cent in Finland (Newman et al., 2024). Many see digital technologies as the principal force that is “dramatically transforming our world” (UN, 2024b: para. 1). According to the United Nations, the danger of mis- and disinformation “cannot be overstated. Social media-enabled hate speech and disinformation can lead to violence and death [...] and endangers democratic institutions and fundamental human rights” (UN, 2023a: 3). Many efforts are underway to govern information ecosystems in ways that are responsive to diverse interests, but how effective are they? What else can be done?

The analysis in this book examines research across the social sciences, humanities, and computer science. Our focus is on research that investigates the causes and consequences of transformations in the news media industry, the development and use of artificial intelligence (AI), and the uses and governance of AI and data within today’s information ecosystems. A commitment to respecting international human rights may guide many responses to mis- and disinformation, but the ways in which Big Tech business models adhere to this commitment and affect the lives of children and adults

online differ, as do the ways that governments are seeking to regulate these companies. Today's information ecosystems are severely testing institutional capacities to sustain the integrity of information, but accurate information is essential if there is to be informed dialogue in the public sphere. If "the right to know is the right to live" (Roy, 2004: para. 1), then securing inclusive, open, tolerant, and respectful online spaces is imperative.

In our assessment of research, we assume that the design, deployment, and uses – both beneficial and harmful – of digital technologies are not dictated *by* technological change; rather, the contribution of these technologies to information ecosystems is understood to be the result of human decisions and actions (Acemoglu & Johnson, 2023). This means we need to examine unequal distributions of power between Big Tech companies, states, and publics, and what this means for efforts to strengthen information integrity and to govern the data which is at the heart of Big Tech business models. We position risks and harms associated with mis- and disinformation as symptoms of complex changes in society *and* as important amplifiers of these changes. As Tay and colleagues (2024: 1) observed, "depending on individual and contextual factors [... misinformation] can be both a symptom and a cause". We use a sociotechnical perspective to broadly inform our analysis of the multidimensionality of information ecosystems.

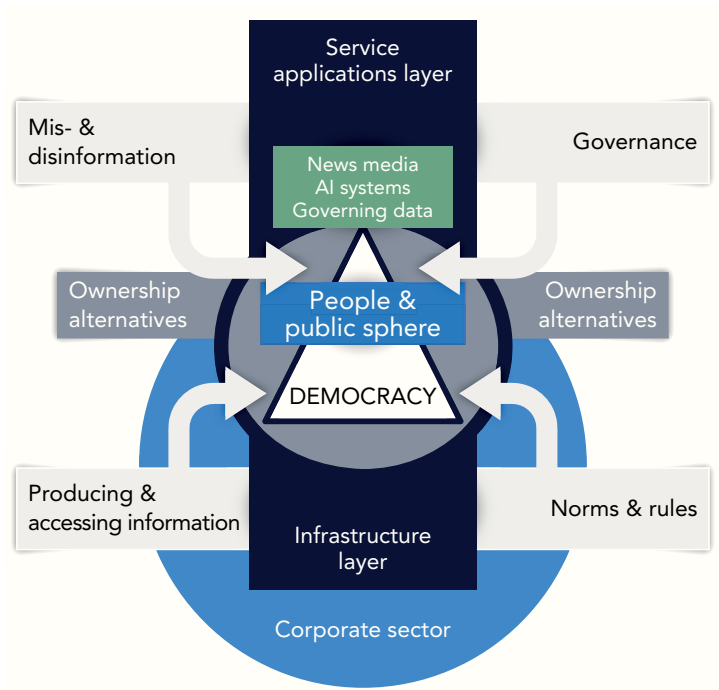
This chapter introduces the concepts of information ecosystems and the public sphere, which are at the centre of our analysis. We then explain why we focus on Big Tech datafication practices and the platformisation of news media, the growing use of AI systems, and the exploitative data practices of Big Tech companies. Information integrity and mis- and disinformation are discussed next, first considering the definitions of these terms and then the challenges in measuring the scale and severity of harms linked to mis- and disinformation. Next, we explain why a discussion of human rights as well as collective interests is key to understanding how information ecosystems could be governed so that they are less troubling for democracy. We then comment on the scope and limitations of this assessment of research and conclude with a short guide to the chapters.

Information ecosystems and the public sphere

A critical analysis of research requires a vocabulary to name the objects and processes we are concerned with. In understanding information ecosystems as comprised of social and material components in a system embracing people, practices, values, and technologies in a particular environment (Nardi & O'Day, 1999), we locate people and the public sphere within a network infrastructure layer and a service layer of information ecosystems. The network infrastructure refers to the hardware and software that supports communication (e.g., data links, routers, switches, compute and data centres, standards, and protocols) and the actors that produce technologies as well as their values and practices.

The service layer refers to services available to users (e.g., cloud, payment, search engines, messaging, app stores, social networking, and e-commerce) and the values and practices of those who design and operate them. Figure 1.1 shows a simple schema of an information ecosystem.

Figure 1.1 Information ecosystems and the public sphere



The private sector plays a major role in deciding how the technological components of information ecosystems are designed and operated, but non-corporate ownership is also of interest in this research assessment. Information ecosystems are in constant flux as power relationships change, and the layers of information ecosystems can be depicted in different ways depending on the country or region (see van Dijck, 2020). The components interact at local, national, regional, and global levels, informed by the social, cultural, political, and economic conditions. When these interactions facilitate mis- and disinformation, they can be an impediment to democracy, and we consider whether decisions about these interactions are, or are not, being based on value judgements that are consistent with democracy. Our analysis is informed by an understanding that the exercise of power involves agency and that it operates on the individual and institutional levels (Jasanoff, 2015; Mansell, 2012; Musiani, 2022; Suchman, 2023). It is also informed by a branch of systems theory that explains the dynamics of changes in information

ecosystems by focusing on networks of humans and non-humans, not on the individual information consumer (Radsch, 2023e).

Research on information ecosystems (or ecologies) appeared in the 1990s in relation to multimedia and, from the early 2000s, in relation to mobile communication and the World Wide Web. From the 2010s, an ecosystem discourse appeared in relation to media convergence and social media as well as digital technology. In recent research, a “healthy information ecosystem” discourse has appeared, defined as a “balanced and well-functioning system of information creation, exchange, flow and utilization” (Radsch, 2023e: 2–3; see also Introne et al., 2024). In the academic literature, the information ecosystems concept is criticised for obfuscating the requirements for democratic governance when it serves as a “means of justification under contemporary neoliberalized orders that typically chafe at modes of public intervention and the language of democratic statecraft” (Gibson et al., 2023: 1442; see also Gillespie, 2010). A “healthy” information ecosystem may be seen as serving the interests of Big Tech companies, shielding them from responsibility for creating problems and serving as “a useful tactic for journalists and policy-makers, who do not recognize or may wish to cloak their own normative positions” (Gibson et al., 2023: 1452; see also Ehrenfeld, 2020). For these reasons, the “health” of an information ecosystem is treated in our analysis as a value judgment, not as indicative of the need to eliminate a pathogen. We also acknowledge the need to draw distinctions between ideal (normative) systems and individual and institutional practice.

The information ecosystems concept may also be said to neglect the historical specificity and role of imaginaries in structuring the public sphere. Some argue that the “public sphere”, instead of the ecosystem concept, should be the focus of research on what influences the fragility of democracy. This suggests a closer focus on whether public deliberation can proceed rationally and on opportunities for deliberative conversations and for the participation of subaltern counter-publics in the public sphere (Dutta & Pal, 2020; Ehrenfeld, 2020; Fraser, 1992; Habermas, 1962/2015, 2022; Mouffe, 1966). The public sphere concept is criticised for its normative assumptions about “critical-rational publics”, its assumptions about liberalism and illiberalism, its Eurocentric emphasis on ideal speech conditions, and its neglect of conflictual political discourse and the role of emotions in public debate (Banaji, 2024; Cammaerts, 2024; Dahlberg, 2014; de Sousa Santos, 2018; Devenney, 2009; Gerbaudo, 2022; Habermas, 2022; Splichal, 2022b; Štětka & Mihelj, 2024a). In our analysis, we use the public sphere concept to centre communicative practices associated with democracy, recognising that “new forms of solidarity, new forms of intimacy, new forms of collective action, and new forms of identification – in short, new forms of being ‘public’ with strangers” (Ehrenfeld, 2020: 308) are needed if democracy is to flourish.

Platformisation, datafication, and democracy

In this research assessment, processes of platformisation and datafication are broadly understood to be undermining democracy. Big Tech companies are using their algorithmic systems to analyse behaviour and to keep people engaged in interactions that generate data and income (Nieborg & Poell, 2018; Plantin et al., 2018; van Dijck et al., 2018a). Datafication refers to the “transformation of social action into quantified data” for real-time tracking and prediction, and it enables companies to transform everyday actions into quantified data (van Dijck, 2014: 198; Mayer-Schönberger & Cukier, 2013). Datafication involves computational methods applied in ways that result in the untransparent manipulation of information and communication flows and the transformation of previously computationally invisible activity into data (Gitelman, 2013; van Dijck, 2014). This platformisation of information means that the economy and multiple spheres of public and private life are influenced by the choices of Big Tech companies (Poell et al., 2019). The companies and many states aim to be leaders in the global economy; for example, the European Commission’s *AI Continent Action Plan* states that “Europe must act with ambition, speed and foresight to shape the future of AI in a way that enhances our competitiveness” as well as democratic values (EC, 2025: 1). The argument in Europe, the United States, China, and beyond is that competitiveness and national (regional) economic growth depend on greater efficiencies in the collection and monetisation of data, and that “technological accelerationism” is good for humanity (Caballero & Monje, 2024).

The technology companies’ business models are implicated in surveillance capitalism or data colonialism; these companies insist that online interaction generates raw or neutral data that belongs to no one (until it is appropriated by them). Building upon decades of research on untransparent offline advertising techniques (McGuigan, 2023), individuals and societies are being comprehensively surveilled for data extraction and products for sale in the digital marketplace. The monopolisation of information generates revenue by converting data (including public data) into private information assets. Even if digital platforms and AI systems open a space for a dynamic between humans and technology that is potentially beneficial, how they operate is decided largely by big companies within legal frameworks put in place by governments. Governance arrangements determine what information “can appear, how it is organized, how it is monetized, what can be removed and why, and what the technical architecture allows and prohibits” (Gillespie, 2010: 359). When someone shows interest in a type of political content, an algorithm is likely to amplify the proportion of similar viewpoints in their feed, narrowing the range of information they see. These practices, combined with the positioning of individual freedom as the enemy of equality and solidarity, are implicated in contemporary social and political instability (Calhoun et al., 2022; see also Bennett & Livingston, 2023; Couldry & Mejias, 2019; Fendji,

2024; Lee & Valenzuela, 2024; Lehdonvirta, 2022; Mejias & Couldry, 2024; Trappel, 2019; van Dijck et al., 2018a; Zuboff, 2019).

Commercial datafication does not fully explain political or economic divisions in society – these “exist before and beyond” these companies’ activities (Tonnies, 1887/1957: 140). Even though online mis- and disinformation are not the only factors that threaten democracy (Aruguete & Calvo, 2023; Zuazo & Aruguete, 2021), when information is inaccurate, harmful, or illegal (e.g., violent or associated with nationalism, ethno-religious bigotry, or misogyny), the risks to individuals and groups multiply. Big Tech companies dominate in providing the components of information ecosystems, but other digital intermediaries such as network operators and web-hosting companies play an important role as well, since they can shut down the Internet or take websites offline unilaterally, especially when they are under pressure from governments (Bradshaw et al., 2021; Bradshaw & DeNardis, 2022).

The public release of generative AI (GenAI) in 2022 has meant that tools for information manipulation are becoming widely available and less costly. AI models, including large language models (LLMs), contain unavoidable biases that stem from data generation and collection processes that are subject to human decisions. People’s own biases influence their online behaviour and interpretations of information, and these biases are exacerbated when AI models are then trained on the data generated by these behaviours. This diminishes the quality of information over time, contributing to decisions that perpetuate inequality and increase vulnerabilities (Corbett-Davies et al., 2017; Dolata et al., 2022; Kop, 2020; Wang et al., 2024). This can negatively influence how citizens understand themselves as political actors, with disproportionately negative effects on marginalised people. Questions are thus raised about how epistemic rights can be protected (Horowitz et al., 2024; Liveriero, 2020). According to Nieminen (2024: 15), “epistemic rights are about knowledge – not only about being informed, but also about being informed truthfully, understanding the relevance of information, and acting on its basis for the benefit of oneself and society as a whole”. To achieve equality in decision-making, it must be guaranteed that accurate information is available to all.

The companies that are marketing advanced digital technologies – including GenAI – as quickly as they can often claim to adhere to safety standards, responsible innovation practices, and to be creating “guard rails” that will reduce the possibility of mis- and disinformation and other kinds of harmful content. However, the pace of these developments can sideline or overcome countervailing power when it is mobilised in regulatory or civil society actions. This is because these companies have created information ecosystems governed by private rule-making embedded in automated technologies – a way of working that contributes to the declining integrity of information within information ecosystems. It is preempting meaningful political deliberation about rights of data ownership, what role data should have in the economy

and public sector, how it should inform bureaucracy, and in what contexts data production should be minimised or prohibited.

This outcome is visible in the transformations in the operations and editorial choices in the news media industry. In principle, the news media can hold the powerful to account and facilitate the free exchange of accurate information, but the quality of public debate suffers when virally circulating information is wrong or inaccurate. The commercial imperatives of algorithmic-driven and opaque advertising mean that political news appears next to sensationalist content, and legacy news media are facing competition from podcasts and bloggers. Market concentration in the legacy media industry, alongside the proliferation of online information flows, is creating a financially unstable news media industry. Datafication and the platformisation of news media mean that news media professionals feel pressured to make their content more attention-grabbing to adapt to digital platform affordances, sometimes sacrificing quality of content (Chadwick, 2017).

Mis- and disinformation are seen as diminishing the quality of news media and public discourse (Bennett & Kneuer, 2023; Jungherr & Schroeder, 2021; Schlesinger, 2020; Wasserman, 2020a), but the contribution of news producers to the public sphere depends on whether pluralism and diversity are encouraged (Hallin & Mancini, 2004, 2012; Neff & Pickard, 2024). In the face of microtargeting during elections, the use of biased AI systems, the rise of deepfakes, the escalation of cyberattacks, and the weaponisation of information, problems facing the news industry are central to the future of information integrity (Briant, 2024; Caulfield et al., 2023; Craig et al., 2023; Forum on Information and Democracy, 2023, 2024a; Holt, 2023; O'Connor, 2022). Research tends to neglect the role of legacy news media and the history of propaganda (Benkler et al., 2018; Bolin & Kunelius, 2023; Hyzen, 2023; Tsfati et al., 2020) when the focus is principally on the impact of technological change. Assessments of the impact of mis- and disinformation – its scale and severity – also need to be informed by the non-technical factors that influence information integrity.

Information integrity and mis- and disinformation

Our aim is to assess what the literature tells us about what changes in information ecosystems and the public sphere mean for communication, for the integrity of information, and for the future of democracy. We understand information integrity as “access to relevant, reliable and accurate information and knowledge” (UN, 2024b: para. 33), and we assume that democracy cannot thrive when information ecosystems are prone to the “substitution of lies for factual truth” (Arendt, 1968: 257). As Arendt (1978: para. 2) observed:

If everybody always lies to you, the consequence is not that you believe the lies, but rather nobody believes anything any longer. [...] And a people that no longer can believe anything cannot make up its mind. It is deprived not only of its capacity to act but also of its capacity to think and to judge. And with such a people you can then do what you please.

We examine research to understand how communicative processes work within information ecosystems and the public sphere. Communication is understood to refer to the exchange of information between individuals or groups using shared concepts and signs, including direct conversations, commercial and public service media, and as mediated by digital platforms. It is through communication that information is gathered and shared, voiced and heard. A well-functioning democracy needs effective communication, which depends on the availability of accurate information. How information influences public opinion and decision-making hinges on fair and open communication within the public sphere, and some information might be helpful in contributing to knowledge that guides behaviour in accordance with social norms and that upholds fundamental human rights. The presence of mis- and disinformation creates the risk that helpful or useful information is crowded or drowned out. This can increase the fragility of democracies and jeopardise human rights protections – the right to freedom of expression, privacy, equality, and justice – and compromise adherence to the rule of law (Advox Team, 2024; Wagner et al., 2025).

Some scholars argue that the scale of mis- and disinformation circulating within information ecosystems is not a significant problem because the causal impact on individuals' attitudes and behaviours is hard to demonstrate, and that this kind of information is a small proportion of the information that many people engage with. For example, there is evidence that only three per cent of active social media accounts are toxic, although these are estimated to produce 33 per cent of that content, and that 74 per cent of all conflicts online start in only one per cent of communities (Robertson et al., 2024). Others argue that mis- and disinformation can “causally and adversely influence people’s beliefs, decisions, and behaviors” (Ecker et al., 2024b: 1).

Defining mis- and disinformation

Disputes about the scale and severity of risks and harms of mis- and disinformation are partly a reflection of inconsistent definitions. Mis- and disinformation as well as hate speech are often depicted as “polluting” information ecosystems (UN, 2023a). Though definitions centre around threats associated with the declining “accuracy, consistency and reliability of information” (UN, 2023a: 5), what is considered to be misinformation or disinformation is shown in Table 1.1, which indicates that an intent to knowingly cause harm is the criterion used to distinguish between these concepts.

Table 1.1 Distinguishing between misinformation and disinformation

	Awareness of falsity	Underlying intent
Disinformation	Aware	Bad
Misinformation	Unaware (inadvertent)	Good / neutral

SOURCE: Bontcheva et al., 2020: 25

In this view of appropriate definitions, disinformation is “information that is false and deliberately created to harm a person, social group, organization or country”; misinformation is “information that is false, but not created with the intention of causing harm”; and malinformation is “information that is based on reality, used to inflict harm on a person, organization or country” (Wardle & Derakhshan, 2017: 15).

The European Union defines disinformation as “verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public, and may cause public harm” (EC, 2018: 3–4). In most cases, the concern about mis- and disinformation is with how “waves of false or misleading content can undermine social cohesion, cast doubt on factual information, and undermine trust in public institutions” (OECD, 2024: 14; see also Altay et al., 2023a; EC, 2020a; EC HLG, 2018; François, 2019; Kapantai et al., 2021; Möller et al., 2020; Ó Fathaigh et al., 2021; Wardle, 2018).

Hate speech is defined in the *United Nations Strategy and Plan of Action on Hate Speech* as that which “attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, color, descent, gender or other identity factor” (UN, 2023a: 5, citing UN, 2019: para. 2). Hate speech is less present in research on mis- and disinformation than might be expected. This may be because hate speech is subject to strong protection of free speech rights in only some jurisdictions (Gillespie, 2020; UN, 2024e), and there is no definition under international human rights law, although “incitement to discrimination, hostility or violence” is prohibited (UN, 1966: Art. 20.2).

Definitions of mis- and disinformation in research vary with respect to the nature and kinds of harm and the level of specificity and granularity that is examined. Whether harm refers to individuals, groups, organisations, or to countries or democratic processes, as well as whether harm must be proven to have occurred, also varies (Wardle & Derakhshan, 2017). The focus may be on state-sponsored campaigns or government, anti-government, or other political propaganda, or it may be on content manifested through individual contributions. Studies may emphasise mis- and disinformation that appear in legacy news media, online news media, or the feeds of online services (Bontcheva et al., 2020). In African regions, in India, and in other countries

in the Global Majority World, definitions are sometimes aligned with those in the Global North, but they are also distinctive, with varying practices regarding sanctions that are applied when information is deemed illegal or harmful (Africa Center for Strategic Studies, 2024; Madan, 2021).

Measuring mis- and disinformation

It is difficult to gauge the scale and severity of the impact of mis- and disinformation accurately. Most assessments are based on one-off studies of a single platform or focus on limited issues. Problems also occur in identifying, gathering, and analysing data that fully reflects people's online experiences. The scale of mis- and disinformation generated and amplified using AI systems is particularly difficult to measure because sources of evidence vary in quality, level of detail, and overall reliability. One review of the growing range of measurement tools and methodologies noted that most are not designed to meet the requirements of regulators when they need data to better understand how to respond to risks and harms in the online world:

The sheer vastness and diversity of online experiences makes meaningful measurement a challenge requiring investment and innovation. The scale and variety of online platforms, and algorithmic personalisation of content, means that there is essentially an infinite number of possible user journeys, making it hard to arrive at both meaningful summary insights as well as fine-grained assessments of particular issues. (Faculty, 2021: 2)

Research on the scale and severity of mis- and disinformation is limited by problems in accessing data. This affects research on revenue flows in the news industry, datafication processes, and decisions in AI systems' development and supply chains. Legal and ethical issues around the collection and analysis of personal and pseudo-anonymised data create barriers to data collection in order to protect individual privacy. Access to government and corporate data is limited, although steps are being taken to address this. For example, according to the European Union's Digital Services Act (DSA), "data access and scrutiny provisions" for "vetted researchers" require access to data of certain platforms and search engines when research concerns the "detection, identification and understanding of system risks in the Union" (EC, 2022c: Art. 40; Forum on Information and Democracy, 2024c). However, the quality of the data provided is controversial, and Big Tech companies have been actively lobbying governments to allow them to prevent data access for research and other non-commercial purposes (Elkin-Koren et al., 2024; Saurwein & Spencer-Smith, 2020; van Drunen & Noroozian, 2024).

Studies of the scale and severity of risks and threats of mis- and disinformation come mainly from research with samples of platform users. These studies use quantitative, survey, or experimental methodologies and, less often, qualitative focus groups or interview-based approaches.

Quantitative studies are typically limited to a few platforms (Facebook, X/Twitter, or YouTube), and they are largely centred on the United States. One study of the diffusion of Facebook posts showed that misinformation diffused via a small number of users and that when content moderation was more aggressive near the time of the 2020 United States election, there were dramatic decreases in the spread and reach of misinformation (González-Bailón et al., 2024). The results of many studies cannot be generalised because they do not take account of conditions – political, social, and cultural – that influence information ecosystems in other parts of the world, and especially in the Global Majority World, where evidence is difficult to obtain (Valenzuela et al., 2024).

There is considerable evidence that mis- and disinformation can lead to negative impacts on public health, the quality of which depends heavily on the information available to those seeking healthcare (Lewandowsky et al., 2022, citing Evanega et al., 2022; see also Gollwitzer et al., 2020; Pennycook et al., 2020). A study in Ireland of more than 13 million posts on 12 online platforms between 2020 and 2023 found that the influence of the far right was growing, with support for white nationalism, antisemitism, and Islamophobia, as well as Holocaust denial and hateful claims about the LGBTQ+ community. Big Tech companies were found to be failing to enforce community guidelines, and the content on smaller platforms was found to be more extreme than on the most-used mainstream platforms (Gallagher et al., 2023). A study of “online hunting grounds in Indonesia” found that the spread of hateful content on TikTok and YouTube was a concern and was seen as receiving little pushback from authorities or the platform companies (Ong & Tapsell, 2022). There is also evidence that when female journalists write reports on mis- or disinformation, digital conspiracy theories, or far-right extremism, attacks on them increase, devaluing their public voice and professional legitimacy (Posetti et al., 2022).

Human rights and governing information ecosystems

Compromises in the adherence to the rule of law occur because democracies are based on a normative order that enables processes of societal self-determination. In democratic orders, the public legitimises complex norms and values, creating a fundamental structure for society that defines how a state and its relations with other actors operate. The exercise of political authority and the distribution of goods and services depend on this ordering, which is coupled with narratives that legitimise and stabilise the normative order (Hamelink, 2023; Kettemann, 2022; Puppis et al., 2024). These narratives are based on information and develop through communication processes. In well-functioning information ecosystems, those who are impacted by decisions taken within that order are assumed to play a role in defining the rules governing which decisions are taken. When the rules for automated systems

and communicative practices are set by actors not perceived as legitimate, decision-making processes become destabilised or corrupted.

Information ecosystem norms and values are subject to international human rights agreements, including the Universal Declaration of Human Rights (UDHR), which states that “everyone has the right to freedom of opinion and expression” and “everyone shall be subject only to such limitations [on rights and freedoms] as are determined by law solely for the purpose of securing due recognition and respect for the rights and freedoms of others” (UN, 1948: Art. 19, 29). In addition, the International Covenant on Civil and Political Rights (ICCPR) asserts that “everyone shall have the right to hold opinions without interference. [...] The exercise of the rights [...] carries with it special duties and responsibilities” (UN, 1966: Art. 19). International human rights law protects the rights and freedoms of the individual, and value judgments on what constitutes respect and responsibility for others can create marked differences in how rights are institutionalised and practised. Embedding human rights, duties, and responsibilities in information ecosystems is a continuous challenge, but new issues are confronted as digital technologies provide novel means of creating and circulating information and speeding up communication processes.

Policymakers are developing frameworks for governing the components of information ecosystems, which means they must find a way to balance national (or regional) goals of companies active in digital markets and the political ambitions of governments, while also securing the rights of the public. On the global level, the United Nations' *Global Digital Compact* calls upon its member states to “commit to respect, protect and promote human rights in the digital space” (UN, 2024b: para. 22) and to uphold international human rights law:

We must urgently counter and address all forms of violence, including sexual and gender-based violence, which occurs through or is amplified by the use of technology, all forms of hate speech and discrimination, misinformation and disinformation, cyberbullying and child sexual exploitation and abuse. We will establish and maintain robust risk mitigation and redress measures that also protect privacy and freedom of expression [... protecting] the rights of the child in the digital space, in line with international human rights law, including the Convention on the Rights of the Child. (UN, 2024b: para. 30)

The compact calls for measures to promote diverse and resilient information ecosystems by strengthening independent and public media; facilitating access to independent, fact-based information; strengthening data governance and multi-stakeholder and broadly-based public engagement; providing human oversight of AI to identify and mitigate risks; and co-developing industry accountability frameworks. These aspirations are echoed in earlier statements by the G7, OECD, The Bletchley Declaration, and UNESCO, and guidance

to encourage enhanced information integrity is set out in United Nations principles (Bletchley, 2023; G7, 2023; OECD, 2022c; UK DSIT; UNESCO, 2023b; UN, 2024d).

The global principles for information ecosystems governance must be translated into institutional practice, which requires robust evidence and a capacity to understand why erosions of principles occur, especially in times of geopolitical tensions and political polarisation (Gillwald, 2024). Approaches to governance intended to achieve equity, redress, and international solidarity differ, and it is essential to understand which publics and whose interests are protected by the governance institutions and practices that are put in place (AI Now Institute, 2024).

The United Nations *Governing AI for Humanity* report acknowledges that “whole parts of the world have been left out of international AI governance conversations” (UN, 2024a: 8). Governance approaches differ not least because in the Global Majority World, some 2.6 billion people are not connected to the Internet (in low-income countries, 20.9% of people use the Internet compared with 90.5% in high-income countries) (ITU, 2024b). Fixed Internet connectivity is either absent or unaffordable for many, although this may be partly compensated for by mobile Internet, though it can be unreliable and costly. Large differences exist in mobile broadband subscriptions per 100 inhabitants (a world average of 90.7 conceals a low-income-country average of 35.3 versus a high-income-country average of 123.4). The world average cost of a mobile data and voice high-consumption basket of services as a percentage of Gross National Income per capita is 4.7 per cent (the low-income country average is 18% and the high-income country average is 0.9%; ITU, 2024a). Disparities within countries, especially between rural and urban areas, are important considerations in understanding how information ecosystems are experienced, as are arrangements for sharing Internet access accounts (Heeks, 2022; Mutsvairo & Ragnedda, 2019; Okolo, 2023; Strover et al., 2024).

In this research assessment, we focus mostly on the experiences of those who are connected to the Internet – more than five billion individual Internet users in 2024 or 67 per cent of the world’s population, with an estimated five billion active social media user identities (not necessarily unique individuals) (Thompson & Kemp, 2024). We do not explore the complexities of digital divides directly, but we do recognise that social, political, cultural, and economic factors influence the production and circulation of mis- and disinformation and that information reaches those who are unconnected through communication in offline social networks and engagement with legacy news. In addition, the economically disadvantaged are often offered Internet service contracts that limit exposure to diverse sources of information. In some countries, there is a legal obligation to participate in political processes even though access to reliable information may depend on the quality of Internet access as well as on social practices, economic policies, and governance, all of

which influence political discourse (Gillwald & Welelakeba, 2024; Hargittai, 2021; Heeks, 2022; Helsper, 2021; Ragnedda & Ruiu, 2020; Robinson et al., 2020; Schaake & Fukuyama, 2023; Trappel, 2019; Warschauer, 2004; Yates & Carmi, 2024).

In the Global Majority World, policymakers often confront decisions taken in the Global North and struggle to govern their information ecosystems in ways that reflect their interests. With Big Tech companies monopolising digital service and data markets, the space for imagining and experimenting with governance alternatives is diminished. Rule-setting by Big Tech companies can be experienced as “behemoths” practising “digital imperialism” (Aaronson & Leblond, 2018; Chen & Gao, 2022), especially when the export of governance models is couched in the language of aid, cooperation, and trade. Within the African Union and in African countries, for example, governance “instruments tend to emulate best practices from other regimes” with unintended consequences when they not “suitable for, or overlook African realities” (Musoni et al., 2024: 15).

Scope and limitations

Our analysis focuses on three main themes with a cross-cutting theme of mis- and disinformation: media, politics, and trust; artificial intelligence, information ecosystems, and democracy; and data governance and democracy. These are outlined in the Appendix. The Appendix also explains the process of creating a large corpus of academic publications, reports, and other materials (3,095 items). These were reviewed resulting in 1,664 citations (plus an additional 51 citations to works published between October 2024 and May 2025) in this book. The corpus is mainly based on academic research, but it benefits from research undertaken or commissioned by civil society organisations and other non-university institutions. Our aim was to map research and policy, synthesise the strengths and weaknesses of research, and highlight new lines of research and potential governance responses. We aimed to assess research in relation to how and why information ecosystems are problematic and for whom, and what is being done, or could be done, to mitigate problems. We were interested in the resilience of governance institutions in the wake of mis- and disinformation and in evidence of their complicity in practices that lead to injustice. We also reviewed evidence on the resilience of individuals in their interactions with online content, but we do not regard individuals themselves as ultimately responsible for accommodating structural injustices (Banaji, 2024).

No assessment of information ecosystems can cover every aspect. Our analysis does not focus directly on research on the unequal material conditions of people’s lives, but it does include work examining how poverty and unequal socioeconomic conditions can lead to exclusions from, and discriminations within, information ecosystems. Our corpus does not encompass the

very substantial field of research on digital divides, but, as above, we do acknowledge huge variations in the availability of meaningful Internet connectivity and restrictions placed on access to information. The balance in Global North and Global Majority World sources in our corpus favours the Global North (see the Appendix). This means our analysis tends to privilege Global North experience (especially the United States and western Europe). Throughout, however, we emphasise that societies experience information ecosystems in distinct ways. Our analysis is also mainly focused on country- or region-level experience, on institutional practices, and on the service layer, although we do examine the network infrastructure layer to some extent. We do not cover the extensive literature on Internet governance, that is, the rules and standards governing the operation of the Internet. Other fields of research – cybersecurity, securitisation, geopolitics and digital sovereignty, the economic geography of digital labour markets, and the (micro)economics of digital markets – make only occasional appearances. Finally, we were inevitably limited by the fact that all research is guided by the questions that research communities set out to examine, the funding available, political pressures on researchers, and access to relevant data.

Conclusion

A global and large-scale critical analysis of state-of-the-art research helps to tell us what we can be reasonably confident about, what is controversial, and what the priorities should be for future research and policy action on information ecosystems. “Scientific knowledge cannot be understood as absolute” (Ecker et al., 2024a: 30), and this absence of certainty creates dilemmas for policymakers when they seek guidance on how to govern information ecosystems in a way that will not abridge human rights (Radsch, 2022). However, the absence of certainty does not mean there are no valid standards for making judgments about how to strengthen information integrity, combat mis- and disinformation, and achieve data justice. Throughout the analysis in the following chapters, we review the literature with the intent of calling attention to what we learned especially from the differences among countries and regions and what those distinctions mean for approaches to governing complicated information ecosystems.

In Chapter 2, the focus is on the news media industry, its structure, and its increasing dependence on digital platforms provided by Big Tech, as well as on the challenges created by declining trust in news media and the increasing weaponisation of information. In Chapter 3, research on how the integrity of information is being influenced by AI systems is examined together with the implications for human rights and democracy. Chapter 4 turns to the way exploitative processes of datafication are entangled with the power of Big Tech companies and the impact of their strategies on decisions about how data is collected and used. In Chapter 5, we examine approaches to

governing network infrastructures, digital platforms, news media, AI systems, data, and mis- and disinformation. We also include an analysis of initiatives to develop media and information literacy and AI literacy. In Chapter 6, we first focus on people's (including policymakers') awareness of mis- and disinformation and the automated content moderation systems they engage with. Measures to combat mis- and disinformation are then examined, including providing information to the public through fact-checking and other means of verifying information, government actions to legislate and regulate Big Tech companies, and platform self-regulatory practices. We also examine alternative data governance frameworks and practices that offer opportunities for resisting exploitative Big Tech datafication strategies and discriminatory uses of AI systems. In Chapter 7, we summarise what our analysis highlights about the strengths and weaknesses of research approaches to concepts and different research strategies. We especially emphasise the distorted picture of information ecosystems that is fostered by the predominance of research from and about the Global North and the urgency of working towards the decolonisation of research. A final section sets out priorities for future research.

News media, information integrity, and the public sphere

Introduction

What is known about the multiple causes and consequences of changes in legacy and online news media, and what can be done to promote strengthened information integrity and a democratic public sphere? In this chapter, we examine the increasing dependence of news organisations on digital platforms provided by Big Tech companies. News media industry market structures and the power relations between news organisations and digital platforms are examined to highlight how the platformisation of news and efforts to monetise news content are creating incentives for the production and circulation of mis- and disinformation. We then examine research on who engages with news and factors that account for people's trust in the news, as well as their perceptions of the trustworthiness of news media organisations. We review studies of changes in journalism standards and practices, including in the wake of growing use of AI systems, and the uses of news and reasons for avoiding news. With information being weaponised by numerous domestic and foreign actors, it is essential to understand how people engage with this information, whether it is a principal cause of polarised publics, and the extent of resilience to harmful or illegal mis- and disinformation.

The challenges facing news media producers and their audiences are examined, drawing upon research that looks at the effects of news media engagement on people's attitudes and behaviours and at studies of the complex factors that influence how diverse information ecosystems are experienced in different political and economic contexts, nationally and regionally (see, e.g., Benequista et al., 2019; Couldry & Curran, 2003; Curran & Hesmondhalgh, 2019; Curran & Park, 1999; Mano & Milton, 2021; Wasserman, 2018).

Defining "news"

In a chapter focusing on the changing environment for news media production, circulation, and consumption, it is necessary to delineate what is considered "news". This is difficult when online content can be produced and circulated by multiple actors. When we discuss "news media", we refer to legacy news

media – television, radio, and offline newspapers – or to coexisting online news media. We also discuss the news media *industry*, which we take to aspire to certain professional norms such as the protection of sources and the goal of objectivity or impartiality. News organisations may be commercial businesses relying on advertising revenues, state-owned public service media (or “public media”), or collectively owned. In some cases, the term “digital journalism” is used to refer to “practices of newsgathering, reporting, textual production and ancillary communication that reflect, respond to, and shape the social, cultural and economic logics of the constantly changing digital media environment” (Burgess & Hurcombe, 2019: 360). We refer to those who may claim to be producing news but who may not share professional news production norms as “actors”.

Some news media organisations may be designated as “alternative media”, and these attract numerous labels: radical, citizens’ media, advocacy journalism, participatory, independent, activist, and grassroots autonomous media. These organisations differ substantially in how they position themselves in relation to mainstream news and in their political orientations (Cushion, 2021). Sometimes defined as “a range of media forms and practices, from radical critical media to independent media, and from grassroots autonomous media to community, citizen and participatory media” (Jeppesen, 2016: 54), these news media forms may position themselves as counter-hegemonic, emphasising non-commercial amateur production with limited financial resources, or in relation to the use of technology for emancipatory purposes (Harlow, 2017).

Legacy and online news media typically share several characteristics, including recording information with digital technologies, news formats intended to engage audiences, and the production of content that can be accessed at any point in time or location. In this context, public service media and collectively owned smaller organisations struggle for prominence on digital platforms. Interactions between news organisations and those who engage with news content shape news media agendas, although not with identical powers of decision-making. These interactions are understood to influence the perceived legitimacy of news media content and the perceived trustworthiness of news outlets. In turn, this relationship is influenced by instrumental and social features resulting from users’ interactions with technology and from the governance arrangements that apply to both legacy and newer news producers (see Evans et al., 2017; Hopkins, 2020; Mazzoli, 2020; Neubaum & Weeks, 2023; Parcu et al., 2022; Ronzhyn et al., 2023).

News media, the public sphere, and democracy

News media organisations play a vital role in meeting the public’s need to be informed about matters of public value. Ideally, they contribute to democracy by fostering community-building, enabling communication flows among members of society, keeping citizens up to date with events, and educating

them (Anderson, 1983; Hanitzsch & Vos, 2018). News media organisations and professional journalists are expected to uphold normative goals embodied in the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights (UN, 1948: Art. 19; UN, 1966: Art. 19) (see also Chapter 1). These rights entail both duties and responsibilities, and the normative goals are “not self-executing”, especially not in countries where democracy is fragile or in authoritarian states (Milton & Mano, 2022: 35; Hamelink, 2023).

In liberal democracies, these normative goals are broadly understood to mean that the news media’s role is to voice the concerns of the public and to hold the powerful to account – “speaking truth to power”. The notion of a public sphere is central to the ideal of democratic deliberation (Habermas, 1962/2015), and an independent news media industry is generally expected to provide factual, accurate, and impartial (or objective) information. This view is challenged when the public sphere is not inclusive, and in his more recent work, Habermas (2022) stressed the importance of distinguishing between the normative conditions that are needed to sustain a democratic polity and the empirical reality of the conditions under which participants struggle to secure the rights to which they are entitled. Inclusive information flows are crucial because they influence the quality of public discourse, the formation of public opinion, and its perceived legitimacy. If information is misleading or wrong, public discussion cannot be fair, and the quality of democratic discourse suffers. The public sphere, historically and today, involves exclusions and marginalisations, which means the ideal of a critical-rational public is often not met (Gerbaudo, 2022; see also Bennett & Kneuer, 2023; Dahlberg, 2014; Devenney, 2009; Hallin & Mancini, 2012; Jungherr & Schroeder, 2021; Rugh, 2000; Schudson, 1978; Schlesinger, 2020; Wasserman, 2020b).

In any given context, multiple public spheres can coexist, and people participate in public life with unequal power, often as counter-publics, especially in the Global Majority World when inclusion is sought from the margins and there is a legacy of colonialism. Thus, the news media does not inform a singular public or operate as the only source of information; there are many sub-audiences to whom news producers can appeal (Dutta & Pal, 2020; Fraser, 1992; Fraser & Nash, 2014). News media may struggle to achieve impartiality, especially in the face of overtly illiberal conditions and constraints on how critical the news media can be of governments in both the Global North and the Global Majority World (Besson, 2011). This means that there are large variations in how closely the ideal of a democratic public sphere is achieved. The responsibilities of the news media are often contested when they conflict with the goals of actors who seek to secure power and privilege (Mukhudwana, 2021; Pintak & Ginges, 2008; Pintak & Nazir, 2013; Repucci & Slipowitz, 2022; Romano, 2013).

The role of the news media and its contribution to the public sphere are strongly influenced by national political conditions and by the production

and circulation of mis- and disinformation. This means that local reporting norms vary, and, in many places, journalist safety is at risk (Conroy-Krutz & Koné, 2022; Chiumbu & Munoriyarwa, 2023). Human rights norms may be universal, but there are many ways of aspiring to these norms, and in the Global Majority World, local practices may differ from those in the Global North (Chasi & Rodny-Gumede, 2022; Wasserman, 2020a). Studies of the news media's role in democracy often present a Eurocentric or Western understanding of how normative goals should be interpreted. Beyond the West, many scholars insist that the rights embodied in international declarations and covenants must be interpreted through the prism of their own cultures. Failures to do so are considered symptomatic of the Eurocentricity of knowledge production. This is especially visible when “particular dominant social, political or economic interests” influence how the news media industry operates (Willems, 2014b: 418).

Journalists, including women and other minority or disadvantaged groups, often face threats, violence, and murder, which are linked to increases in the production and circulation of mis- and disinformation (including hate speech). Since 1993, some 1,701 journalists have been reported as killed, with 50 per cent of the deaths occurring outside conflict zones (UNESCO, 2024). In 2020, 73 per cent of 625 female-identifying journalists said they had experienced some form of online violence (UNESCO, 2022d). In relation to women and their rights, “misinformation, disinformation and defamation are real and pervasive threats [...] and women tend to be targeted more frequently than men” (Gallagher, 2023: 58). In the Global Majority World (and elsewhere),

the act of gaining voice and attempting to have influence in the public sphere, the act of asserting their own humanity and right to exist qua human beings, takes place during ongoing deprivation, campaigns of disinformation, police brutality and/or military atrocity. (Banaji, 2024: 13)

For this reason, it is important to not lose sight of the structures of power that influence how the news media operate or the material conditions of people's lives when we focus on how to combat mis- and disinformation and on efforts to build a robust public sphere.

News media industry structure and platform dependency

Reporters Without Border's *World Press Freedom Index* reports that the economic indicator for press freedom is at an “unprecedented” low because of “ownership concentration, pressure from advertisers and financial backers, and public aid that is restricted, absent or allocated in an opaque manner” (RSF, 2025: para. 2). Research on news media market concentration and the news media's structural dependence on Big Tech platforms reveals why many legacy and online news media organisations are facing a crisis that threatens their sustainability.

The history of asymmetric global and regional news media markets and news flows between the Global North and Global Majority World was studied long before the Internet. The nonaligned movement (NAM) – a group of 120 mostly low- and middle-income countries active from the 1950s and remaining active – elected not to align themselves with, or against, any major power bloc. This group debated the hegemony of news organisations mainly in the Global North and sought to establish a New World Information and Communication Order, partly to address the news media’s role in promoting propaganda (Mansell & Nordenstreng, 2006; Thussu, 2022; Vincent & Nordenstreng, 2015). There is a long-standing tradition of research on the way news media have participated in propaganda in the Global North, especially during World War II (Benkler, 2020; Lasswell, 1971; Rantanen, 2024). Today, the news industry is in crisis in many parts of the world as news organisations confront declining advertising revenues due to Big Tech digital platform domination of the advertising market (Recuero, 2024). The way news media organisations are responding to this crisis differs around the world, even as journalists and their organisations benefit from hosting news on digital platforms that reach new online audiences.

The Big Tech companies have encouraged the “platformisation” of the news industry with “the rise of digital technologies, in a neoliberal, political, and economic climate” and the ubiquity of digital platforms in people’s lives (Plantin et al., 2018: 298; Garcoa Ramirez, 2021). This has created pressure on the news industry to change its operations. The companies providing digital platforms have become akin to publishers, and even editors, because news production relies on their algorithms, advertising markets, data, and content-moderation standards. Captured by the platform companies, many news outlets have come to depend on financial arrangements established by the Big Tech companies, although the platform owners argue that they benefit news organisations through their public relations campaigns and informal relationships (Greene, 2018; Nechushtai, 2018; Nieborg & Poell, 2018; Papaevangelou, 2023; Radsch, 2023b).

Structural asymmetries are present because of the power of both dominant news media producers – and the Big Tech companies – to control how audiences are exposed to news. This power is exercised as a result of business models geared to profiting from the monetisation of data generated by the audience’s online interactions and encouraged by algorithm-driven news personalisation systems that promote mis- and disinformation (Mansell & Steinmueller, 2020; Wasko et al., 2011; Winseck, 2022). Hosting mis- and disinformation that appears on websites is a lucrative business estimated to have generated 2.6 billion US dollars in advertising revenue in 2021 worldwide, with the United States accounting for more than half. Meta generated at least 30 million US dollars in ad revenue from content it had removed from its own platforms for engaging in coordinated inauthentic behaviour (Elliott, 2022; Skibinski, 2021).

Studies of structural power in the news media industry are conducted at country level, and often comparatively, although differences between types of news producers that operate sub-nationally are also examined (Nieborg et al., 2024; Nielsen & Fletcher, 2023; Nielsen & Ganter, 2022). The market structures of the privately owned news industry create economic and political incentives for the production and circulation of news content (Rohn et al., 2024), and these structures are visible in ownership conditions, the extent to which news organisations are politically independent, and in the relationships between digital platform companies and the news organisations. These power relations and structures influence how the news media create “the terrain for other actors to contest power” (Freedman, 2014: 324; see also Thussu, 2022) and how concentrated “media power” is seen as anti-democratic:

Concentrated media power [...] is antidemocratic both because it hands definitional, analytical, and interpretive power to unelected organizations and because it undermines the ability of citizens to acquire and exchange the information and ideas necessary to make informed decisions about public life. It is also dangerous, because it distorts the logic of the media industries themselves, transforming them from vehicles of symbolic interaction to increasingly significant engines of capital accumulation. (Freedman, 2014: 327)

News media organisation dependency is due to the Big Tech companies’ ability to aggregate end users, which magnifies network effects (Montero & Finger, 2021; Nieborg & Poell, 2018; Nielsen & Ganter, 2022) and results in their market dominance in the Global North (Bannerman, 2022; Moore & Tambini, 2018, 2021) and in the Global Majority World (Mabweazara et al., 2020; Mabweazara & Pearson, 2024). Big Tech companies’ control over the advertising market gives them the power to dictate financial terms to news organisations (Bell et al., 2017; Garcia Ramirez, 2021; Nielsen & Ganter, 2022; Radsch, 2023c). The platforms’ leverage comes from their position as intermediaries between news media content producers and their publics, and their capacity to deprioritise news (Kristensen & Hartley, 2023; Nielsen & Ganter, 2022; Poell et al., 2023). News media advertising on the platforms does attract traffic for some news publishers, but this rarely translates into economic sustainability for their businesses (Nieborg & Poell, 2018). News media organisations have attempted to monetise their content by tailoring news to the platforms’ affordances (i.e., their technical characteristics and rules of operation) to boost user engagement and advertising revenue (Bell et al., 2017). The news organisations’ reliance on ad tech metrics (e.g., clicks/impressions as a news performance indicator) may have created a more competitive newsroom culture, but it also has increased management surveillance (Petre, 2021). Table 2.1 highlights tensions in the relationships between some of the largest digital platforms and news media organisations.

Table 2.1 Tensions between digital platforms and news media organisations

Platform	Relationship between digital platforms and news media organisations
X (formerly Twitter)	Elon Musk's takeover of Twitter was followed by platform changes that were perceived as problematic or unfavourable for the journalism profession. In January 2024, X suspended an unknown number of prominent accounts that were critical of the Israeli government, blaming the spam algorithms. In October 2023, X stopped showing headlines in previews to improve the aesthetics of the iOS app. The measure of excluding titles from previews of links might have been meant to discourage users from sharing third-party content from media organisations. Focusing on the French media sector in a six-month period after Musk's takeover, it was found that journalists started to question the broader legitimacy of social media as a journalistic tool, but engaged in "strategic disconnection" instead of abandoning the platform.
Facebook (Meta)	Research on the "Facebook problem" includes critiques of technology and design; amplification of misinformation; abuses of market power; monopolisation of the digital advertising market; undermining financial support for journalism; encouraging sensationalist journalism and clickbait; and limitations of self-regulation. Small tweaks to the News Feed's algorithm can have a profound impact on the visibility of news content.
Google	There has been criticism of the monopolisation of online advertising markets. Google's news services tend to steer revenue toward the largest publishers. Google has developed products that enable off-site publishing in new formats, such as Google Accelerated Mobile Pages. This implies a loss of control over channels of communication and increased dependence on platforms as news intermediaries. Risks include losing control over editorial identity (search algorithms shape the way users interact with news content) and access to data (more detailed analytics are available on-site than off-site).
YouTube (Google)	There is concern about changes in YouTube's policies regarding the demonetisation, delisting, and removal of videos, with implications beyond the performance of individual videos.
TikTok	There are concerns about the virality of content directed at young people, the decline of legacy news as gatekeepers, and how its algorithm exploits children's vulnerability and distributes racist and sexist content (in addition to geopolitical security concerns between the United States and China).
Social media platforms in general	There is criticism of dominant social media platforms and their failure to moderate harmful content at key moments (e.g., Gamergate, a misogynistic online harassment campaign; the Rohingya genocide in Myanmar from 2016; the 2016 United States presidential election; the Christchurch New Zealand Mosque shooting 2019; or Alex Jones's promotion of conspiracy theories and the rise of online revenge porn). There are problems with automated moderation and overstated claims of success; difficulty accounting for context, subtlety, sarcasm, and subcultural meaning; and insensitivity to the use of duplicate content in different contexts, such as terrorist propaganda reposted in a journalistic context.

SOURCE: Claesson, 2023; Germain, 2024; Gillespie, 2020; Nielsen & Ganter, 2022; Notley et al., 2020; Peters, 2023; Pickard, 2020c; Poell et al., 2023; Van Natta et al., 2023

In the early 2010s, Germany and Spain legislated in support of the news media in response to Google's Search and News products with mixed results. Australia and Canada designed reforms in response to the unequal bargaining power of the digital platform owners and news media publishers (the News Media Bargaining Code and the Online News Act, respectively), with other countries considering similar legislation, for example, India, New Zealand, the United Kingdom, and the United States. Some news organisations have benefitted from direct payments by platform companies, for example, Google or Meta, as compensation for hosting news content (Hermida, 2023; Marshall, 2023; Nielsen & Ganter, 2022). The platforms derive substantial economic value from featuring news on their sites, although compensation arrangements can lead to the largest news media organisations benefiting disproportionately. Sceptics argue that these measures are centred around disputed claims about the economic value of news, and the long-term relationship between the news media and platforms remains contested (Flew, 2023; Flew et al., 2023; Flew & Martin, 2022; Flew & Stepnik, 2024). Disputes occur about the scale of compensation, especially when it is argued that the digital platforms do not "free ride" on the news media and that payments by platforms to the publishers can inhibit innovation (Lesh, 2023). It has been estimated that the platforms in the United States would owe news publishers between 11.9 and 13.9 billion US dollars annually if an agreement were to be reached on compensation (Mateen et al., 2023). Clashes over payments and how to value news have led on some occasions to the deplatforming of news. Meta, for example, threatened to remove news content in Australia and Canada, and took action to do so when an agreement was not reached (Meese & Hurcombe, 2021).

Despite efforts to try to sustain independent, high-quality, and trustworthy news, newspaper sales are declining globally. In both the Global North and Global Majority World, there are moves to empower smaller news organisations to bargain collectively with Big Tech companies as local news organisations in the Global North have been hard hit as they moved their content online. Competition for audiences is reducing local news stories to little more than clickbait in some countries (Tomaz & Trappel, 2022), and news deserts have been reported in Europe, the United States, and other regions as news organisations close at local and sometimes regional levels.

Public service media, which is also sometimes called "public media" (including organisations permitted to attract advertising), are rarely economically sustainable without subsidies, concessions, protections, or a combination, that involve direct government financial support, license revenue, technical assistance, and collaborative strategic programming and advocacy (Radsch, 2022). And news media organisation dependence on the platforms is especially severe in low-income countries where press freedom is limited or non-existent (BBC Media Action, 2021; Garcia Ramirez, 2021); Nielsen & Ganter, 2022; Schot, 2020).

It is important to differentiate between the changes in the news media industry in regions and countries and between legacy and newer online media. For example, in countries in Latin America, where financing often comes from philanthropists, legacy news media are pressured to focus reporting on the interests of their funders (Labio-Bernal & Romero-Domínguez, 2022). Some argue that too great a focus on the sustainability of legacy news media organisations distracts attention from the historical and ongoing marginalisation of certain audiences from the public sphere (Usher, 2023; Verza et al., 2024).

The challenges facing news media organisations are compounded by their problems in accessing audience data from the platforms (Meese & Hurcombe, 2021; Nieborg & Poell, 2018; Nielsen & Ganter, 2022). News organisations benefit from the ability to monitor audience interest, and this requires verification data. There are difficulties in accessing accurate data and in interpreting the partial data journalists receive, which is biased to favour the platforms (Dommett, 2023). In the United States, the Federal Trade Commission (FTC, 2024) has been investigating how data accuracy impacts the market share and the potential for news sales growth, especially with the expansion by AI companies into media markets, which is creating further pressure on news provider finances. Instant online news production also undermines news verification processes and the ability of journalists to fact-check mis- and disinformation, with the result that legacy news media struggle to maintain their audiences and the credibility of their news (see Baron, 2002; Carson & Gibbons, 2023; Dierickx et al., 2023b; Himma-Kadakas & Ojamets, 2022; Ross Arguedas et al., 2022b; Wahl-Jorgensen & Carlson, 2021).

To generate revenues, some news media organisations have their own online news distribution sites, some are hosting paid-for events, and others have turned to native advertising (ads with the look and feel of content they appear with) hosted at their own news sites (Meese & Hurcombe, 2021). Tax credits, direct subsidies, philanthropic funding, not-for-profit models, and subscription or membership models are being tried. Efforts to influence the production, distribution, and monetisation of news content may include using news organisations' own apps (mainly viable for the largest providers) and setting up paywalls or membership programmes, creating their own real-time advertising marketing capability, or launching cooperative news organisations (Grover & Baik, 2024; MacKenzie et al., 2023; Marshall, 2023; Poell et al., 2023). Another means of financially supporting news media is taxing digital advertising, which has been proposed (Radsch, 2022), with strong criticism, by the United States government and, especially since the Trump Administration in 2025, treated as an unfair trade barrier (Higgs, 2025). An additional problem for news organisations is that some Big Tech owners of platforms are becoming less interested in hosting online news as they seek new sources of revenue growth from the integration of AI tools into their platforms. This is likely to create new challenges for news publishers when they try to build interest in their own online sites and to attract readers.

None of these financing strategies is individually likely to fully address the news media industry's funding gaps. One challenge is demonstrating that a public subsidy or a tax is substantial enough to compensate for the historical cross-subsidy of news content by advertisers to news organisations. With advertisers attracted to the platforms instead of directly to news organisations, it is argued that the cost burden falls on the news-consuming audience. It must then be demonstrated that news has value for the health of the public sphere and democracy, beyond the measures used to value news economically. There is substantial literature on efforts to sustain the financial independence of news organisations and the impacts on democratic rights and freedoms when these efforts fail (see Altay et al., 2023a; Baines & Elliott, 2020; Epstein, 2020; Fallis, 2015; François, 2019; Kapantai et al., 2021; Milton & Mano, 2022; Ó Fathaigh et al., 2021; OHCHR, 2021: para. 9–15; Pielemeier, 2020; Willems, 2014a) and specifically on financing independent news (see Borges & Reviglio, 2024; Brogi & Sjøvaag, 2023; Henningsen & Krčál, 2024; Kennis, 2024; Mansell & Steinmueller, 2020; Mutsvairo, 2019; Pickard, 2020c, 2024; Rohn et al., 2024).

These developments are contributing to the production and circulation of mis- and disinformation driven by platform profit motives and by the exploitation of the affordances of platforms by professional persuaders (Bakir & McStay, 2018). The actors producing and circulating this information take advantage of the algorithm-driven ad tech market that facilitates “digital deceit” to amplify content (Ghosh & Scott, 2018; Pielemeier, 2020). This generates significant revenue for the platforms, for those advertisers willing to have their content appear alongside this content, and for individual “influencers” (e.g., celebrities or creators of fake accounts and information).

In summary, market concentration in the news media industry and in the platform market distorts news organisations' capacity to contribute to a robust public sphere and undermines democracy (Bimber & Gil de Zúñiga, 2020; Vaidhyanathan, 2022). As Sevignani and colleagues (2025: 806) have argued, “tech giants do not simply capture ‘the news’ by means of acquiring news outlets. [...] Big Tech, by influencing our information environment, shapes our lifeworld consequently privatizing the public sphere”. With news organisations facing political and financial pressures, and with many dependent on digital platforms to circulate their news, this leads to questions about their independence (Wasserman, 2018). A result of these market structural features is that news media organisations are prevented from fulfilling their democratic function, media pluralism is reduced, and this contributes to perceptions of the news media industry as untrustworthy and to variations in individuals' trust in the news.

Trust in news media

History is replete with examples of partisan (and political party-funded) news media that have voiced the concerns of segmented audiences rather than the

public. This history varies from country to country and with whether public service media have served the needs of the public in an impartial way. As discussed above, the structure of news media markets and their dependence on digital platforms provided by Big Tech companies are important contributors to whether the news media are trusted and whether new media organisations are considered trustworthy. “Trust” may be defined in various ways, but it is generally associated with,

[an] individual’s willingness to be vulnerable to media objects, based on the expectation that they will perform a) satisfactorily for the individual and/or b) according to the dominant norms and values in society (i.e., democratic media functions). (Fawzi et al., 2021: 156; see also Strömbäck et al., 2020)

Research on trust in news media often focuses on whether media power and influence are concentrated in ways that limit the diversity and plurality of news content. Research consistently finds that some Western countries are experiencing declining trust in legacy news media, but not in all countries or at the same rate. Some research focuses on individual trust in news media, while other work focuses on relationships, recognising that “trusting is not a matter of blind deference, but of placing – or refusing – trust with good judgement [... and] we need social and political institutions that allow us to judge where to place our trust” (O’Neill, 2002: 7; see also Blöbaum, 2016; Frislich & Humprecht, 2021). It is broadly agreed, however, that when the news media are considered untrustworthy by the public, this contributes to the undermining of democracy.

Evidence of declining individual trust in news

Evidence of diminishing trust in news obtained via platforms is varied (Ross Arguedas et al., 2022c; van Dijck et al., 2018a). In the Global North, as news consumption habits have steadily moved from legacy news media to online sources and social media platforms (Ofcom, 2023c), research comparing trust in news that appears on digital platforms (Facebook, YouTube, Google) in Brazil, India, the United Kingdom, and the United States has found that trust depends on the platform, the country, the audiences, and the kinds of news, and that these sources are less trusted than legacy news media (Mont’Alverne et al., 2022). Brazil has seen one of the steepest declines in trust in the news, dropping from 62 to 43 per cent between 2015 and 2024, with the far right playing a key role in distrust of legacy media. In the West, some countries do not seem to be affected by declining levels of trust, for example, in Denmark, trust is relatively stable, at around 57 per cent. In South Korea, a relatively low level of trust is reported to have risen from 22 to 31 per cent between 2016 and 2024. There are also countries where trust is already high and still increasing: In Kenya, it went up from 50 to 64 per cent between 2020 and

2024, and in Thailand, from 50 to 54 per cent between 2021 and 2024. The United States has seen a strong and asymmetrical decline in trust in news between left-wing and right-wing voters, with a similar pattern in some countries beyond the West (Newman et al., 2024; see also Hanitzsch et al., 2018; Newman et al., 2022; Strömbäck et al., 2020).

Interviews with news workers in Brazil, India, the United Kingdom, and the United States indicate how platforms hamper audiences' ability to distinguish between credible and non-credible news sources (Ross Arguedas et al., 2022c). Across 47 markets and six continents, survey results show that concern about online "fake" news reached an average of 59 per cent, although this varied by country and region (e.g., South Africa, 81%; the US, 72%, and the UK, 70%; and Germany, 42% Norway, 45%, and India, 58%) (Newman et al., 2024: 17). Another thing to consider is that not everyone accesses the news, and trust in the media may also be associated with a specific media element – such as a person (a journalist, an expert), a source (e.g., *The Financial Times*, Fox News), a type (television, radio, press) – or it may apply generally to "the media" (Skovsgaard & Andersen, 2020). The results of surveys of media trust are questioned by some scholars who argue that it is unclear whether survey respondents understand what journalism standards are, or should be, when they answer survey questions (Bernardi & de Moraes, 2021; Bhat & Chadha, 2020; Christofolletti & Becker, 2023). Research on trust in the media tends to focus variously on overall trust in the news media, on news media as a public institution, on media organisations and their ownership, on media types (television vs. radio), specific outlets, media coverage, and trust in journalists (Fawzi et al., 2021). Trust in "media" is also often examined interchangeably with trust in the news media, in legacy media, and sometimes in online media, which leads to ambiguity in interpreting whether and why trust in news media is changing (Schrantz et al., 2018).

Research using large scale surveys confronts the problem of ambiguities around definitions of trust and news media (Christofolletti & Becker, 2023). These studies reveal associations between trust and individual factors (sociodemographic, political, and social attitudes) based on aggregate data (e.g., at country level) or on individual-level data. Most aim to measure overall trust in news media, which tends to address trust in journalism as an institution, and many studies are based on self-reports that may not be indicative of how people behave. Research using qualitative methods is needed to provide a deeper insight into why people trust or distrust the news.

Models and concepts of media effects

A variety of theories underpin research on media trust. Some research on audience engagement with the news seeks to identify how exposure, including to mis- and disinformation, directly causes changes in attitudes and behaviour by isolating news media impacts from other factors (Anderson, 2021; Klapper, 1960; Lasswell, 1971; McQuail, 2010). The hypodermic

needle model (sometimes known as the inoculation model) of media effects suggests that information will trigger a similar reaction in everyone exposed to it, regardless of people's characteristics (Bineham, 1988). This model was initially developed to understand the effects of government propaganda in the era of mass media, and it grants little or no agency to people's ability to interpret information. A two-step flow model of communication was later developed to add context, positioning opinion leaders as playing a role in mediating between news media and their audiences (Katz, 1957). A selective exposure model proposed that people choose which news media to engage with based on their preexisting views, assuming more limited news media effects (Stroud, 2017). Other research focused on how the media shapes attitudes, suggesting that the media has a cultivation role – that is, audiences tend to view the world as it is depicted in the media (Gerbner et al., 1980). Agenda-setting and framing theories inform studies of the capacity of the media to set an agenda and to influence people's selection of topics that matter to them (Goffman, 1974; McCombs & Shaw, 1972; Valenzuela et al., 2023). These models of media effects are influential, and they benefit from new methods for measuring effects of information that circulates on social media platforms (Choi et al., 2020; Scott et al., 2022).

Approaches to the study of media effects are also informed by theories from behavioural economics. Here, the focus is on cognition and on the effects of nudging people away from mis- and disinformation based on understanding affective and cognitive responses. Insights into cognition are used to provide cues to encourage people to change their online behaviour and attend to the accuracy of information, and this work is largely based on experimental studies. This approach assumes a limited-attention utility model derived from the economic and psychological analysis of how choices are influenced by people's preexisting preferences (Pennycook & Rand, 2022). This research finds, for example, that average exposure to mis- and disinformation is not as high as is sometimes claimed, and that social media are not the primary cause of broader social problems, such as polarisation. Exposure to false and inflammatory content has been found to be concentrated within fringe groups with a high motivation to seek out this information (OECD, 2022b; Pennycook & Rand, 2022). This line of research also emphasises that research is needed on exposure to media content among extremists and fringe groups, and on efforts to limit demand for this kind of information by curtailing political elites and legacy media that spread mis- and disinformation (Budak et al., 2024)

Studies that examine individuals' responses to news media find a variety of associations between individual characteristics and reported trust in news media. These studies rely on quantitative data collected at the individual level, finding, for example, that in Germany, people with higher levels of interpersonal trust (the propensity to think that others will not harm them) report higher levels of trust in the news media (Jackob, 2012; Tsfaty & Ariely, 2014). In the

United States, research finds that those with higher levels of political cynicism have less trust in the media (Frieden, 2014; Pinkleton et al., 2012).

The sociodemographic and age factors associated with trust in news media are inconsistent across countries. In some countries, men are less trusting, while the opposite is found in, for example, Israel, and no association between trust and gender was found in the United States (Schranz et al., 2018; Toff et al., 2021a; Tsfatı & Ariely, 2014). In Brazil, India, the United Kingdom, and the United States, less educated people were found less likely to trust news media (Toff et al., 2021a). In contrast, a study across 44 countries found a slight decrease in trust with each additional year of schooling, and longitudinal research showed that those with a higher education degree were slightly less likely to trust the news (Hanitzsch et al., 2018; Tsfatı & Ariely, 2014). There is also evidence that diaspora communities tend to make greater use of non-mainstream media, while long-term residents make much less use of it (Trauthig, 2024).

Ideology and partisanship can influence trust in news media. This research is dominated by evidence from the United States, finding, for example, that Republicans are associated with significantly lower levels of trust in the media (Toff et al., 2021a; Verma et al., 2018), and that trust in local and national news organisations has declined more rapidly for Republicans than for Democrats (Eddy, 2024), which is seen as a form of asymmetric propaganda (Benkler, 2020). In other countries, trust is found to be more closely associated with attitudes towards extremism and populism than with left–right commitments. Some studies find that those who situate themselves in a more extreme ideological position are less likely to trust the media; others – those with stronger populist views – tend to trust the news media less (Hanitzsch et al., 2018; Stroud & Lee, 2013; Suiter & Fletcher, 2020). Studies show that extreme ideology is positively associated with beliefs in conspiracy theories in Sweden and, in the United States, it is also a predictor of lower trust in legacy news media, although those engaged with conspiracy theories may still have an interest in news (Krouwel et al., 2017; McKernan et al., 2023). Interest in and knowledge about politics are found to influence trust in media, and cross-country studies show that interest in politics is positively associated with trust in the news (Hanitzsch et al., 2018; Tsfatı & Ariely, 2014).

The news media does not necessarily exacerbate mis- and disinformation problems. A study in Brazil, India, and the United Kingdom investigated the effect of news and platform use on awareness of, and belief in, Covid-19 misinformation. News consumption weakened the acquisition of false beliefs depending on the information access mode (online or offline) and the news outlet type (Altay et al., 2023b). A study of news exposure in Mexico during its 2021 midterm elections examining the relationship between frequency of news exposure on social media platforms and beliefs in political mis- and disinformation yielded results consistent with a minimal media effects

paradigm, although platforms using visual communication and emphasising strong network ties were more influential (Valenzuela et al., 2024).

The reasons people distrust the news also vary. Perceived convergence between the interests of journalists and politicians or businesses and a belief that the powerful push an agenda is found to be strong among young people and those with lower incomes (Newman & Fletcher, 2017). An interview-based study suggested that suspicion about the neutrality of news media leads media users to doubt the media, with concern about who to trust and what to believe (Toff & Nielsen, 2018). How the media industry reports news also influences trust, with perceived accuracy, impartiality, expertise, and integrity shaping the perceived quality of news and the level of trust (Kantar Media, 2016). People's subjective perceptions of accuracy also influence trust. For example, when people with direct experience of an event believe there is a difference between what they witnessed happening and how it was reported, their trust in the media is impacted (Livio & Cohen, 2018). Research also shows that rumours that go viral are often more influential than the credibility of a source of information or its factuality; sharing such information is found to be motivated less by the accuracy of information than by "partisan support, community sentiment, emotional contagion and a taste for the sensational or bizarre" (Rodríguez-Ferrándiz, 2023: 15). However, the operationalisation of measures of affect or emotion in these studies relies on inconsistent definitions (Altay et al., 2023a, citing Rogers, 2020; Bakker & Lelkes, 2024; Wardle, 2023).

The role played by news media in circulating mis- and disinformation (formerly propaganda) long predates the Internet. The affordances of platformised media change the distribution of power between the news media and its audiences, as indicated above. This raises questions about the role of algorithms in shaping public beliefs and behaviours (Benkler, 2020; Guess et al., 2023a). As the information engaged with on digital platforms infiltrates people's lives, this is seen as threatening democracy (Ross Arguedas et al., 2022a; Tucker et al., 2018). The research in this section seeks to understand this mainly by measuring the effects of audience exposure to mis- or disinformation as a way to explain the effects of news media on people's attitudes and behaviours.

Relational models of media engagement

Other research in the sociocultural tradition examines reciprocal relationships between the content provided by the news media, the roles of changing technologies and the broader political, social, cultural, and economic contexts in which news media operate (Schünemann, 2022). The audience research tradition, for example, focuses on how audiences interpret media content, examining how people's lives are mediated by their engagements with online and offline information (Couldry & Hepp, 2016; Silverstone, 2007). Here, it is assumed that audiences have sufficient agency to interpret the news and to

do so in ways that are conditioned by their contexts. Using both quantitative and qualitative methods, audience reception studies seek to understand how audiences and the media co-produce information and cultures. Studies focus on the uses and gratifications that audiences experience when they engage in news selection (Livingstone, 1998), and engagement (or non-engagement) with legacy and online media is considered important for people's – and especially young people's – abilities to make sense of the world. Those who engage online may be depicted as living "inside" media, with research focusing on how teenagers construct identities through online interaction (Deuze, 2014), or on why audiences choose their preferred media diets (boyd, 2014). The audience research tradition emphasises that audiences use technologies in unexpected ways and that they engage actively with news media content.

It was concluded several decades ago that,

despite the volume of research, the debate about media effects – whether it can be shown empirically that the specific mass media messages, typically those transmitted by television, have specific, often detrimental effects, on the audiences who are exposed to them – *remains unresolved* [emphasis added]. (Livingstone, 1996: 306)

Nevertheless, the search for effects of mis- and disinformation circulated by the news media and other actors continues to be informed by "effects theories" with the aim of discovering ways to mitigate harm. It is argued that research on media effects today is experiencing a pendulum swing back to the assumptions of the hypodermic needle media effects model, which are misleading in an era of datafication (Ong & Das, 2020).

The "effects" tradition persists, despite findings of weak media effects. For example, in a study in Chile, there was a weak relationship between mis- and disinformation and media scepticism. Initial beliefs about factually dubious information were negatively correlated with levels of trust in the news media in this panel study. Although lower trust in the media was related to higher levels of mis- and disinformation, the strength of this association weakened over time. There was no evidence of a positive feedback loop – the reverse spiral model – between mis- and disinformation and media scepticism, with initial beliefs about factually dubious information negatively correlated with levels of trust in the news. This work called into question research asserting "the corrosive effects of misinformation on democratic regimes" (Valenzuela et al., 2022: 368).

It is also important to acknowledge that declining trust in news media may partly be a response to rising scepticism about political and other institutions. As media historian Schudson (2022: 150) has noted,

trust in institutions is salutary for democracy only to a point. The decline in trust in most institutions that public polling has documented since the 1960s was a decline from what was arguably much too unquestioning a level of trust.

Increasing levels of distrust may be desirable if this is associated with critical thinking, and declining trust in news media might be a sign of a more intellectually active public, although such distrust can also be associated with nativist and racist sentiments. It is also important for journalists to engage with questions about what constitutes truth, accuracy, and facts to avoid being branded as the “lying press” as documented in, for example, South Africa (Wasserman, 2020a).

Changing journalism standards and news media practices

Research on the changing role of news media in information ecosystems focuses on journalism standards and news media practices to assess transformations in the industry. Substantial resources have been historically needed to produce and disseminate news, and legacy news media functioned as gatekeepers, selecting what they deemed to be important (Shoemaker & Vos, 2009). When this power was accompanied by adherence to widely promoted standards of reporting (e.g., accuracy, impartiality), this was considered a positive contribution to democracy and information integrity. In the current news industry environment, the news media often excessively amplifies certain narratives. In some countries (e.g., Brazil, France, Italy, Spain, and the United States), the news media amplifies far-right narratives when they report the discourse of far-right populist figures on issues such as immigration, foreign affairs, the environment, or gender discrimination (Pérez-Curiel et al., 2021). During the 2016 United States presidential election, for instance, research suggests that the propagation of mis- and disinformation took advantage of structural weaknesses in the country’s media institutions (Benkler et al., 2018).

Challenges in realising journalistic values associated with the trustworthiness of news media organisations raise concerns about declining editorial control in news media. News media organisations face the challenges of maintaining editorial authority independent of platforms and governments, maintaining high-quality news standards, and delivering in-depth and diverse content (Eichler, 2023; Hartley et al., 2023; Nielsen & Ganter, 2022; van Dijck et al., 2018a, b). During World War I and II, some news organisations sought to distance themselves from state propaganda, claiming to emphasise facts and objectivity, especially in the United States (Schudson, 2022; Tuchman, 1972). However, in other circumstances, such as the ending of Apartheid in South Africa, the news media faced difficult trade-offs between encouraging the new democratic government and criticising its actions (Wasserman, 2020b). Although the spread of digital platforms may have led to the platformisation of news and more diverse voices in the public sphere, the Big Tech owners of digital platforms are now seen as operating as a “fifth estate” due to their gatekeeping role (Dutton, 2023), and this raises questions about journalistic values (Tandoc & Vos, 2016).

As journalism co-evolves with social media platforms, journalism practice “not only symptomatically reflects, but also pragmatically adapts to and influences the changing media environment” (Burgess & Hurcombe, 2019: 360). Journalism is not being “transformed by being digital”, but it is a practice that embodies the use of digital technologies (Duffy & Ang, 2019: 378; Zelizer, 2019). A meta-analysis of studies published between 2013 and 2018 indicates a shifting focus from research on how digital technologies impact journalism to how journalism reflects and impacts society (Steensen et al., 2019). This work acknowledges that the discourses, practices, and logics of journalism are reshaping the cultures, technologies, and products of news media.

As an agent of change, “digital journalism” is seen as influencing the status and role of digital platforms (e.g., the legitimisation of X/Twitter as a global newsroom). It is also seen as triggering changes in technical processes and practices (e.g., the growing popularity of news content on Facebook led Meta to acknowledge its editorial and curatorial role and responsibility for the content it fosters, which then led to changes in its algorithms and to greater efforts to signal contested news, harmful content, and mis- and disinformation on its platforms) (Burgess & Hurcombe, 2019; d’Haenens et al., 2022; Erkmén, 2023). The result is that what accuracy and “truth” mean for journalists is understood differently around the world, and procedures for achieving accurate reporting are challenged as digital technologies contribute to changes in journalism practice.

In the digitalised news environment, news organisations risk losing editorial control, which can diminish their credibility and lead to perceptions of news bias. A competitive journalism culture within newsrooms, fuelled by scrutiny of performance metrics and managerial surveillance, intensifies pressures on journalists and can deprioritise investigative journalism. There is evidence, however, of positive outcomes when journalists take advantage of digital services; for example, the use of WhatsApp in Rwanda has helped journalists to extend their coverage and educate each other through debate about their practices (McIntyre & Sobel, 2019). These developments also mean that alternative media organisations and journalists may be co-opted by the far left and positioned as criticising commercial values, while right-wing media is associated with alleged ideological partisanship (Ihlebak et al., 2022). Both offer counterpoints to a dearth of diverse viewpoints, yet alternative news media is said to engage in “one-sided and ideologically motivated campaign journalism” (Ihlebak et al., 2022: 1269). These news media do play a role in fostering dialogue, enabling marginalised voices to be heard and challenging the status quo, even if some are involved in circulating exclusionary narratives that may contribute to audience polarisation (Benkler et al., 2018; Siapera, 2023). Unfortunately, most studies of news diversity exclude alternative media, even when it operates as an influential competitor to legacy media.

News media use, news avoidance, and resilience

Numerous factors influence people's media use, whether they try to avoid the news, and whether they are likely to be resilient to mis- and disinformation. Where news media are diverse and (relatively) free from state coercion, a high-choice news media environment presents opportunities to study which audiences consume what type of news and whether their consumption patterns explain political participation (Chadwick, 2017; Prior, 2005). People have different media repertoires, that is, varying degrees of engagement with media sources and types of content such as entertainment, political information, and regional or national news (Castro et al., 2022). Age may be a strong predictor of media-use repertoires, with consistent evidence that older people tend to use more legacy media (e.g., watching television more than other age groups) and younger people getting their news more from social media (Castro et al. 2022; Kim, 2016; Strömbäck et al., 2018; Taneja et al., 2012). The less educated are more likely to access news from television. People who listen to podcasts tend to be more educated, and there is evidence that short videos are becoming a more common source of news, especially for younger people, although this varies by country (Aalberg et al., 2013; Newman et al., 2024). More educated men are more likely to consume news from legacy media (Castro et al., 2022; Strömbäck et al., 2018).

Varied media repertoires are associated with political participation. In the United States, a robust media system is historically associated with a public that consumes hard news (Schudson, 1978). Soft news and the use of social media tend to be regarded as less noble but are important in shaping people's engagement in politics (Castro et al., 2022; Reinemann et al., 2012). Despite inconclusive results, incidental news exposure is found to lead to reflections on politics that can increase people's knowledge. In a world where a humourist can live-cast a conversation with an actor about politics, the distinction between hard and soft news is fragile.

Some people tend to actively avoid certain sources of information. This is problematic in the presence of polarisation and partisanship, especially if news selectivity leads to more extreme political positions (Buturoiu et al., 2023). Selective news exposure is influenced by several factors, such as confidence in one's judgments and political knowledge, or the degree of belonging to a homogeneous social group (Dubois & Blank, 2018; Metzger et al., 2020). While both political knowledge and interest are predictors of news usage, knowledge may be a stronger predictor of whether people are more likely to seek out news stories rather than avoid them (Lecheler & de Vreese, 2017). Affective or emotional engagement plays an important role in people's news usage or avoidance (Corbu et al., 2021; Zhu et al., 2024). Smaller qualitative studies of everyday news use (as demonstrated in the case of Kenya; Tully, 2022) can reveal the importance of social and cultural dynamics that influence motivations to engage with the news and to share false information on social media and chat apps.

News media avoidance

Studies of news avoidance – people who voluntarily or involuntarily consume very little or no news at all – indicate that this occurs to varying degrees around the world (Skovsgaard & Andersen, 2020; Toff et al., 2023). A study of 46 countries (with data from 2015–2022) showed that the number of people claiming to not participate in any news increased by 19 per cent, with this pattern present in most countries and for most types of news. In some countries, participation was flat (e.g., Austria, Ireland, France, Japan, the Netherlands, and Switzerland), but it was increasing in, for example, Colombia, India, Indonesia, Nigeria, and Peru (Altay et al., 2025). However, variations in how news avoidance is defined mean that research results are not easily compared (Bos et al., 2016; Castro et al., 2022; Strömbäck et al., 2018).



One study found that “people who selectively avoid news consume almost as much news as those who do not” (Palmer et al., 2023: 697), although this involves a mix of deliberate choices and socially conditioned preferences. Another study found that 39 per cent of survey respondents said that they sometimes or often avoid the news. An increase over the previous year, with significant increases in Brazil, Finland, Germany, and Spain (Newman et al., 2024). In France, another survey found that 94 per cent of those aged 15 and older reported an interest in information and that they stayed informed daily. (Arcom, 2024)



Numerous factors are associated with higher levels of news avoidance, including significantly lower consumption or higher avoidance of news by women (Toff & Kalogeropoulos, 2020; Toff & Palmer, 2019), and a higher likelihood to avoid the news by young people (Toff & Kalogeropoulos, 2020). The more educated access more news than the less educated in Norway (Karlsen et al., 2020), and a study in Sweden showed that political interest plays an increasing role in the consumption of news over time – political interests were found to be a determinant of news avoidance, with those with lower interest in politics being more likely to avoid the news (Edgerly, 2022; Strömbäck & Shehata, 2019). Those with a weaker understanding of the news media ecosystem or who tend to trust the media less are more likely to avoid the news (Boukes & Vliegenthart, 2017; Edgerly, 2022; Toff & Kalogeropoulos, 2020). People may also avoid the news when they perceive it to be too pessimistic, for example, avoidance increased during the first

year of the Covid-19 pandemic (Aharoni et al., 2021; Newman et al., 2024; Suiter & Fletcher, 2020; Villi et al., 2022). News avoidance may at times be a strategy to protect one's mental health or to avoid information overload (as found in Argentina, Finland, Israel, Japan, and the United States) (Aharoni et al., 2021; Villi et al., 2022). News avoidance can also be attributed to a coping mechanism or to a form of protection against negativity and constant stimulation (Ytre-Arne & Moe, 2021; Suiter & Fletcher, 2020). People may avoid the news because it is irrelevant to them, they believe the news is not trustworthy, or it is too commercial (Aharoni et al., 2021; Edgerly, 2022). In Argentina, people were found to avoid the news because they regarded its information ecosystem as corrupt, while in Japan, they were more likely to try to avoid controversy and disagreement (Villi et al., 2022). The few studies that look at structural factors find that greater press and political freedom and stability are negatively correlated with news avoidance (Toff & Kalogeropoulos, 2020). Overall, there is no definitive answer as to how much news people should consume, but news avoidance is problematic when it isolates people from daily political discussions and decision-making.

The weaponisation of online information

In the late 1960s, Hannah Arendt discussed whether it is always necessary to tell the truth, distinguishing between factual truths (facts, events) and rational truths (e.g., mathematical, scientific, and philosophical truths) in political debates in plural societies. She explored the disturbing consequences of denying, mystifying, or replacing truths with the opinions of political actors (Arendt, 1968). In today's "post-truth" era, authoritative figures centre political communications around the strategic denial of verifiable facts (Giusti & Piras, 2021; Lockie, 2017; Merenda, 2021). In this way, information is weaponised and contributes to democratic fragility. This is especially so when minorities are singled out via social media to receive divisive and manipulative content (Freelon et al., 2022; Freelon & Wells, 2020; Park et al., 2023).

The weaponisation of information in political contexts, including elections, is a major concern. In 2024, more than 80 countries and some three billion people were set to vote in regional or national elections, although the metric used to count elections does present measurement challenges (Harbath, 2023). The political use of social media and data to target communications directly at followers in unethical (and sometimes illegal) ways to influence election outcomes is central to debates about the harms associated with mis- and disinformation (Elishar-Malka et al., 2020). Politicians increasingly seem to be able to lie without negative consequences, and actors – political elites or digital-platform owners – increasingly claim hegemony over what counts as truthful interpretations of reality (Eady et al., 2024; Hofmann, 2024).

Cambridge Analytica, the political campaign company that operated from 2013 to 2018, sparked outrage as one of the first "information operations" actors to interfere with democratic processes on a grand scale

by microtargeting individual voters and spreading disinformation. They were found to have undertaken illegal data gathering in both the United Kingdom and the United States due to their use of some 5,000 data points on voters, which were secured without user consent via This is Your Digital Life, an app hosted on Facebook. The company's tools for targeting voters were used to discredit its clients' political opposition in numerous countries (Briant, 2021; Dowling, 2022).

Mis- and disinformation can be advantageous to political figures whose supporters share this content (e.g., Donald Trump and QAnon conspiracy theorists). The Trump Administration and Fox News facilitated and co-produced persistent mis- and disinformation during the Covid-19 crisis (Jeppesen et al., 2022; Yang & Bennett, 2022; see also Santini et al., 2018), and research shows that Trump's supporters have influenced the dynamics of top "fake news" spreaders (Boulianne et al., 2022; Bovet & Makse, 2019; Pérez-Curiel et al., 2021).

Analysis of X/Twitter news activity suggests that fake and extremely biased news have distinct amplification mechanisms compared to centre- and left-leaning news. In Brazil, a study found that mis- and disinformation tended to circulate more on political pages and groups aligned with the far right and former Brazilian President Jair Bolsonaro, and on religious and conspiracy theory pages and groups and alternative (hyper-partisan) media, whereas fact-checked news circulated more on leftist pages and groups (Recuero et al., 2022). Another study documented the spread of conspiracy narratives in Brazil about George Soros, providing evidence of cross-platform dissemination (Santini et al., 2020, 2022). However, in the case of electoral misinformation during the 2022 Brazilian presidential election, it was found that professionally produced news from legacy news organisations played a key role in curbing misinformation. Despite this information spreading on digital platforms, there were either no or very small effects between the use of platforms to source news and beliefs in electoral misinformation (Mont'Alverne et al., 2024).

Mis- and disinformation practices are consistently associated with far-right political movements and politicians in Brazil, Germany, Portugal, and the United Kingdom, and studies of migrant-related mis- and disinformation in the European Union reveal a mix of state-driven activities linked to Russia and domestic far-right actors (see Baptista & Gradim, 2022; Buarque & Zavershinskaia, 2022; Chavalarias, 2024; Daniels, 2018; Freelon et al., 2022; Recuero et al., 2020; Szakacs & Bogнар, 2021; Wojcowski, 2022).



The use of legacy media to weaponise information in autocracies varies by social and political context. Right-wing activists spread their messages by manipulating legacy media and working strategically with partisan

media, but there is less research on the magnitude and character of left-wing mis- and disinformation activities. Illiberal political leaders adopt mis- and disinformation as a tool for gaining support and reducing resistance without resorting to terror – by securing formal ownership of the media or informal control, as in the case of information autocrats (e.g., Lee Kuan Yew in Singapore, Augusto Pinochet in Chile, Vladimir Putin in Russia, Alberto Fujimori in Peru, and Mahathir Mohamad in Malaysia), and when there is convergence between parts of the media and far-right political parties (e.g., the far-right party Vox in Spain). In the Middle East, the Arab news media is subject to persistent repression by authoritarian governments, with evidence of the media pushing mis- and disinformation. Russia and China deploy state-owned media outlets such as Russia Today (RT) and China Central Television (CCTV) (Douai, 2019; Freelon & Wells, 2020; Guriev & Treisman, 2019; Labio-Bernal & Manzano-Zambruno, 2023).



However, in Russia, large-scale surveys of citizens' abilities to distinguish between true and false news stories showed that their capacity was comparable to those living in other political contexts. Participants who mostly consumed pro-regime state media made less accurate evaluations than those who mostly consumed independent media. Government supporters were substantially more susceptible to pro-regime misinformation than those who were critical of the regime. Both pro-regime and regime critics were more likely to reject stories incompatible with their beliefs, and true critical stories were rated as false about half the time, suggesting “that the reporting of independent media is often not a threat to authoritarian leaders” (Shirikov, 2024: 61).

In certain African countries, mis- and disinformation are connected to colonial legacies of mis- and disinformation and propaganda rather than to the rise of the far right (Mudde, 2019). In countries such as Kenya, Nigeria, and Zimbabwe, the spread of false information is especially complex when it is polarised and ethnically charged (Mare et al., 2019). And in some African countries, the entire political spectrum, as well as legacy and social media, participate in mis- and disinformation (Lunga & Mthembu, 2019; Ncube, 2019; Wasserman, 2020a). While mis- and disinformation or “fake news” is often treated as a novel scholarly topic today, false news as a phenomenon in Africa and the Middle East pre-dates the era of online news. Journalists have always had to learn to treat journalism as a contested area, vulnerable to manipulation by governments and powerful social elites (Mutsvairo & Bebawi, 2019), although governments now have new opportunities to restrict freedom of expression on social media. State-sponsored mis- and disinformation campaigns are common in multiple political systems (la Cour, 2020). For example, tools, capacities, strategies, and resources for propaganda

have been identified in 81 countries, with private firms engaged in campaigns and practices of harassment against fact-checkers and those reporting on information operations (Bradshaw et al., 2021). Many illiberal leaders are preserving a democratic façade while controlling the information space – acting as informational autocrats (Guriev & Treisman, 2019).

Russian influence operations and information warfare promoting anti-Western propaganda are used to undermine trust in NATO, the European Union, and national governments by interfering in elections and undermining democratic processes (Akimenko & Giles, 2020; Alieva et al., 2022; Beskow & Carley, 2020; Lemke & Habegger, 2022; Morkūnas, 2023; Robbins, 2020; Zhang et al., 2021). Russian propaganda has pushed mis- and disinformation with the aim of winning a significant share of media audiences in countries such as Argentina, Bolivia, Colombia, and Mexico (Chaguaceda et al., 2023). An analysis of the activities of troll factories under the control of the Russian Internet Research Agency from 2017 to 2019 in the United States found that the scale of the Internet Research Agency's troll factories was “industrial – mass produced from a system of interchangeable parts, where each class of part fulfilled a specialized function” (Linville & Warren, 2020: 463). This highlights how, rather than posting tweets in support of one party, Internet Research Agency trolls tweet divisive messages, sometimes targeting mainstream Republicans and, at other times, mainstream Democrats – tactics consistent with sowing mistrust and doubt in the election process (Linville & Warren, 2019). Research during the 2020 United States presidential election examined the promotion by pro-Kremlin media (the channel RT) on Facebook, and how content curation algorithms affected its distribution (Kuznetsova & Makhortykh, 2023). The Facebook News Feed algorithm (which Facebook describes as aiming to expose users to reputable information in a Meta blog post; Brown & Levin, 2020) was found to enable pro-Kremlin media to propagate these messages, indicating that Facebook's content curation is vulnerable to manipulation of “likes” to enhance message flows.

Actors, especially in Russia and China, have been connected to mis- and disinformation campaigns in both the Global North and Global Majority World (Sleibi, 2023; Znojek, 2020). China wages a state-sponsored mis- and disinformation campaign against a United States-led international system to suppress internal and external criticism, amplify its prestige, and favourably influence foreign policy actors, and to support its military strategy (e.g., on the issue of Taiwan) (Cheng, 2016; Curtis, 2021; Hung & Hung, 2022).

Discourse using concepts such as “fake news” is invoked in competitions for power and used to discredit, attack, and delegitimise political opponents (Farkas & Schou, 2018). Legitimate news media are targeted by “fake news” labelling (Tong et al., 2020), reducing the perceived credibility of authentic media content, although research suggests that this may not affect people's policy preferences (Hameleers & Marquart, 2023). The weaponisation of information in these ways has coincided with the explosion of technologies

that make mis- or disinformation part of a flux of overabundant information (Bargaoanu & Radu, 2018). Although search engines play a role in promoting mis- and disinformation, much attention focuses on the role of social media in the creation, distribution, and monetisation of this online content because it sometimes outperforms legacy media as a source of news (Aïmeur et al., 2023; Newman et al., 2023, Wakefield, 2016). In the United States, consumption of legacy media sources was associated with more accurate beliefs about health-related topics and consumption of non-partisan liberal media instead of partisan conservative media, and with a smaller inclination to access “fake news” websites (Jamieson & Albarracin, 2020; Guess et al., 2019). Research on the large social media platforms (Facebook, X/Twitter, Instagram, TikTok), video-sharing platforms (YouTube), and private messaging apps (WhatsApp, Telegram) is summarised in Table 2.2.

Table 2.2 Platform roles in the weaponisation of information (selected countries)

Country	Example	Platform
Argentina (49) Bolivia (27) Colombia (78) Ecuador (65) Peru (28) Spain (127)	Channels of distribution across Spain and five Latin American countries: Facebook was the most-used network to disseminate mis- and disinformation (32.9%), followed by hoaxes disseminated in two or more networks (31.9%), WhatsApp (21%), Twitter (5.7%), e-mail or SMS (5.4%), and YouTube (3.0%)	Multiple social media platforms; bibliometric study of 371 examples of mis- and disinformation about the Covid-19 pandemic (March–May 2020)
Australia (2) England (1) Spain (1) United States (8)	Weaponised crowdfunding by actors to amplify and sustain the spread of their grievances. Crowdfunding platforms offer comprehensive tools that facilitate easy sharing and propagation of campaign messages across various platforms	GoFundMe, a set of election fraud and 5G-themed campaigns on the crowdfunding platform GoFundMe
Canada	Prevalence of misinformation surrounding Covid-19 on Twitter, compared with Canadian news media: social media exposure was associated with more misperceptions and less social distancing compliance	X/Twitter; all articles published on 19 Canadian news sites
India	WhatsApp as a tool for political communication used by political parties in India: for mobilisation, coordination, and reaching out to voters; political propaganda and disinformation were pushed on WhatsApp in the form of “news”	WhatsApp
United States	Evidence from the 2016 presidential election on the virality of political fake news: posts favouring Trump were shared 30 million times on Facebook, while those favouring Clinton were shared 8 million times	Facebook/Meta
United States	Fuelling civil disobedience in democracy: WhatsApp news is negatively associated with political knowledge and positively with illegal protest	WhatsApp

Continued

Table 2.2 Continued

Country	Example	Platform
United States	False news stories (2006–2017) diffused significantly further, faster, deeper, and more broadly than the “truth”. This effect was stronger for political news than for other topics	X/Twitter
United States	Amplifying climate mis- and disinformation: showing that posts linking to content from 10 “superpolluter” publishers, totalling 186 million followers on Facebook, accounted for up to 69 per cent of Facebook interactions with climate denial content	Facebook/Meta; a sample of 6,983 climate denial articles were published between 12 October 2020 and 1 October 2021
Zimbabwe	Digital propaganda “battles”, where political gladiators used mis- and disinformation, hate speech, and mudslinging as weapons	X/Twitter
n.a.	YouTube facilitates access to problematic content (sometimes with mixed results); nine of the studies demonstrated support for the creation of filter bubble effects	YouTube’s systematic review of 23 studies published between 2013 and 2021

SOURCE: Allcott & Gentzkow, 2017; Center for Countering Digital Hate, 2021; Chibwe, 2020; Elmer & Ward-Kimola, 2023; Gutiérrez-Coba et al., 2020; Farooq, 2018; Gil de Zúñiga & Goyanes, 2023); Vosoughi et al., 2018; Bridgman et al., 2020

NOTE: Numbers in parentheses after country names indicate number of incidents covered by cited sources.

Sponsored content and the absence of transparent political advertising rules help to weaponise information. The amplification of mis- and disinformation uses the same tools that are the backbone of online advertising (e.g., precision advertising, algorithmic advertising, data-driven behavioural segmentation, psychographics profiling, computational profiling, and computational persuasion) (Bargaoanu & Radu, 2018; Cano-Orón et al., 2021; Szczepkowski & Szczepkowski, 2021). These developments are examined in multiple studies aiming to establish whether and to what extent (under what conditions) exposure to mis- and disinformation is causing changes in public opinion and leading to polarisation.

Mis- and disinformation, political polarisation, and public opinion

Research on the relationships between mis- and disinformation and filter bubbles or echo chambers and polarised public opinion yields ambiguous, or at least contested, results, and it does not clarify how the news media are implicated due to the varied theories and concepts being used. A filter bubble has been defined as “a unique universe of information for each of us [...] which fundamentally alters the way we encounter ideas and information” enabled by algorithmic prediction engines (Pariser, 2011: 9; Sunstein, 2007). An echo chamber is understood as “a bounded, enclosed media space that has the potential to both magnify the messages delivered within it and insulate them

from rebuttal” (Jamieson & Cappella, 2008: 76). Research focuses on whether engagement with social media makes people angry or affectively polarised; creates echo chambers; amplifies posts that are emotional, inflammatory, or false; and leads to increases in the probability of violence. Many studies are experimental or quasi-experimental and are undertaken in the United States or other Western countries. Research effort in this area intensified from 2014, when social media platforms – Facebook and Twitter (now X) – added “like” and retweet buttons to their sites, and whether social media increases susceptibility to online mis- and disinformation and leads to political polarisation has become a major area of inquiry (Sultan et al., 2024).

Some studies show how exposure to content on social media increases affective or emotional polarisation, while others find limited and asymmetrical effects of social media use on attitudes towards people with diverging views. A study in France, the United Kingdom, and the United States found no evidence that online social media “explain support for right-wing populist candidates and parties”, but it did find that “offline discussion with those who are similar in race, ethnicity and class positively correlates with support for populist candidates and parties in the United Kingdom and France” (Boulianne et al., 2020: 683). A study in the United States found a “substantial amount of overlap (51%) in the ideological distributions of accounts followed by users on opposite ends of the political spectrum” (Eady et al., 2019: 1). A study on the role of social media platforms in contributing to radicalisation and violent extremism found no effect of echo chambers (Gunton, 2022). Another claimed that research overestimates the impact of digital technologies in explaining social and political developments (Talamanca & Arfini, 2022). Again, in the United States, it was concluded that AI tools (LLMs, in this case GPT-3) in the hands of those launching foreign covert propaganda campaigns can be highly persuasive, as measured by people’s agreement with claims (Goldstein et al., 2024). Yet, a study of claims about filter bubbles was challenged in research on public opinion in the United States, which found that social media use had led to less polarisation as judged by partisanship (in this case, vaccine hesitancy), while use of legacy media made people more polarised (Jones-Jang & Chung, 2024).

Studies of single platforms at a single point in time have indicated that in 2021, Twitter’s personalisation algorithm was amplifying tweets in Canada, France, Germany, Japan, Spain, the United Kingdom, and the United States, but did not “support the hypothesis that algorithmic personalisation amplifies extreme ideologies more than mainstream political voices” (Huszár et al., 2022: 4). It did show that the political right experiences higher amplification compared with the political left overall, but the study was unable to identify precise causal mechanisms for the variations among countries.



A team of researchers collaborated with Meta to investigate questions about social media effects on politics using large-scale experiments during the 2020 United States presidential election. One study looked at the effects of echo chambers on polarisation, finding that reducing exposure to like-minded content did not reduce polarisation, although it did decrease exposure to uncivil language and increase exposure to cross-cutting sources (Nyhan et al., 2023). A second study investigated whether Facebook enables ideological segregation in political news consumption, finding that conservatives were more segregated than liberals, and that disinformation circulated mostly in an isolated conservative space. This study found that “ideological segregation is high and increases as we shift from potential exposure to actual exposure to engagement” (González-Bailón et al., 2023: 392). A third study examined the effect of Facebook’s news feed algorithm by sorting posts chronologically, finding chronological filtering did not affect levels of issue polarisation and affective polarisation, although it could increase exposure to untrustworthy content and content from moderate voices, as well as decrease exposure to uncivil content and the time spent on the platform (Guess et al., 2023a). A fourth study looked at the effects of reshared content, finding that removing reshares from the platform could reduce exposure to untrustworthy content, but this did not affect political polarisation (Guess et al., 2023b).



These studies indicate that personalisation systems and interaction with like-minded content can influence consumption of content from untrustworthy sources and exposure to incivility, but they do not demonstrate clear effects of social media on polarisation. This research was conducted over a relatively short time, focused on one country (where polarisation has increased in recent years), and on a single platform. Steps were undertaken to ensure the reliability of the findings and to limit Meta’s influence on the results, but reliance on one platform company’s data can potentially bias the results (2020 Election Research Project, 2020). The reliability of findings was questioned when, for example, it was revealed that impacts of mis- and disinformation may have been influenced by a temporary change in a news algorithm so that it appeared to feed largely reliable sources of trustworthy news to users in contrast to the less rigorous algorithm that was normally used. Bagchi and colleagues (2024) argued that this change was not considered, but Meta, in this case, insisted that it had informed the researchers of the change (Guess et al., 2023b).

These and other studies demonstrate the difficulty in pinning down causal relationships between mis- and disinformation, the news media, the role of algorithms, and platformisation in polarisation and political participation (Ecker et al., 2024; Robertson et al., 2024). This is partly explained by problems of access to platform data, overreliance on controlled experiments, and the scarcity of field studies that seek to confirm hypotheses about the causal effects of mis- and disinformation (Forum on Information and Democracy, 2024c). Studies on polarisation are also resource-intensive, and researchers must collect data through platform application programming interfaces (APIs) that typically allow limited access. The entanglement of content with personalisation systems and social relations mean it is difficult to specifically identify which factors contribute to attitudes towards mis- and disinformation and to political participation.

Research in some countries in the Global Majority World points to positive features of online filter bubbles and echo chambers, which are found to provide a degree of respite from targeted attacks on marginalised groups. This is shown to help protect marginalised groups – for example, feminists, LGBTQ+ populations, those with disabilities, religious groups, or political dissidents – by providing a safe space and possibilities for avoiding political or social repression. For vulnerable populations and disadvantaged or marginalised groups, the appearance of polarised groups communicating in filter bubbles can yield safe spaces to express opinions, and well-conceived algorithms have the potential to enable people to express their ideas and identities without fear of punishment (Erickson, 2024; Toff et al., 2021b).

There is little doubt that differences in news media and information use, experiences of harassment and abuse, invasive data collection, and propagation of mis- and disinformation are implicated in “[fanning] the flames of hatred and division in society” (Bennett & Livingston, 2020: 20). However, research suggests that exposure to like-minded political content is only one of several causes of polarisation of public opinion. Some argue that “politically partisan online news echo chambers are generally small – much smaller than is typically assumed in public and policy debate” (Ross Arguedas et al., 2022a: 17), the claim being that research on the negative impacts of mis- and disinformation exaggerates the harms (Altay & Acerbi, 2023; Allcott & Gentzkow, 2017; Karpf, 2020; McGonagle et al., 2019). In an interview (12 February 2024), Rasmus Kleis Nielsen, then director of the Reuters Institute for the Study of Journalism and professor of political communication at the University of Oxford, suggested that empirical evidence on polarisation indicates that mis- and disinformation and social media algorithms contribute to a small extent. While there is little evidence that most people are influenced by the mis- and disinformation they encounter online, elite cues matter more in terms of impact and influence, and the long-term effects are not well understood, according to Kleis Nielsen. In another interview (13 February 2024), Natalia Aruguete, researcher at the National Scientific and Technical Research Council

and professor at the National University of Quilmes in Argentina, observed that the causal relationship may be that polarisation in society generally is itself a cause of people believing mis- and disinformation. And in a third interview (27 February 2024), Eugenia Mitchelstein, associate professor and chair in the Department of Social Sciences, University of San Andrés and co-director at the Center for the Study of Media and Society in Argentina, pointed out that the availability of platforms has allowed people who could not successfully enter into mainstream media to gain popular recognition and cover a different side of the news (see also Boczkowski & Mitchelstein, 2021, 2022; Mitchelstein & Boczkowski, 2023).

Cross-disciplinary and longitudinal research in this area is relatively scarce (Obreja, 2023). The negative effects of mis- and disinformation on political outcomes and democracy may be assumed at the outset of a study, and in some work, the impacts on truth or trust in expertise, institutions, and the news media are alluded to, but the roles of mis- and disinformation are unclear or unspecified (Allcott & Gentzkow, 2017; Eady et al., 2023; Guess et al., 2023a; Kapantai et al., 2021; LSE Truth, Trust & Technology Commission, 2018; Nyhan et al., 2023; van der Linden, 2023, citing Vosoughi et al., 2018).

Filter bubbles and echo chambers are “not phenomena purely related to algorithms and what information they present, but to how people react to and interact with information” (Talamanca & Arfini, 2022: 19), and this depends on the locale. Research that considers individual agency, market structural conditions (financial and business models) in the news media and platform industries, and the political ideologies that feed polarisation is needed if polarisation phenomena are to be understood (Pickard, 2020a). Nevertheless, mis- and disinformation, and especially hate “in the space in which we debate publicly is one of the main ways of weakening democratic institutions [translated]” (Aruguete & Calvo, 2023: abstract). While platform “algorithms segregate and personalize [...] they cannot on their own, explain entity divisions [translated]”, and polarisation in the public sphere “exists before and beyond [translated]” (Zuazo & Aruguete, 2021: abstract) the algorithm-driven personalisation of news content (Aruguete & Calvo, 2023). Even if vulnerability to mis- and disinformation and its impact on public opinion varies by context, there is no doubt that powerful actors do jeopardise “free and open opinion formation as well as promote the dispersal of communicative power” (Seipp et al., 2023b: 20). Research on the effects of filter bubbles and echo chambers on polarisation needs to acknowledge that what is “unproblematic for one individual can have fatal consequences for another, and which consequences are negative and which are positive is always contestable” (Geiß et al., 2021: 683).

Studies of rumours and information online and their role in political campaigns – even when their accuracy is uncertain – find that repeat spreaders can disproportionately influence public opinion, although there is also evidence of perverse effects of efforts to raise awareness around

deepfakes being associated with distrust in legitimate information (Guess et al., 2023a; Kennedy et al., 2022; Twomey et al., 2023; Weismueller et al., 2024). Identity propaganda aimed at amplifying historic differences and perpetuating hegemonic power structures can influence public opinion through its use of othering narratives and its influence on attitudes and behavioural norms (Reddi et al., 2023). Explanations for some differences in results might become clearer if research takes into account a wider range of contextual factors, including power relationships (Martínez-Costa et al., 2023). Research also finds that the cost of reaching people with mis- or disinformation, not the cost of creating it, is a bottleneck for those intent on distributing mis- and disinformation, and that the evidence for the effectiveness of microtargeting through personalisation is, in any case, limited (Simon et al., 2023).

There is less research (in our corpus) on the role of mis- and disinformation in polarisation and in shaping public opinion in non-Western countries. As in the Global North, there is little consensus in the research community on how to define what counts as fake or false news, and it is often the state that determines what is considered misleading information in autocratic and democratic countries alike. It is also clear that the ontological and epistemic implications of using language like information “disorder” or “pollution” can be symptomatic of ethnocentrism that privileges a Western view of how information should be generated and assimilated in liberal democracies. In the Global Majority World, it is even more important to focus on who is doing the labelling of mis- and disinformation, what is being labelled, and how it is labelled (Banaji & Bhat, 2022; Banaji et al., 2019; Harsin, 2024).

Strengthening trust and resilience to mis- and disinformation

Whether declining trust in news media is perceived as a crisis depends on the country. The emphasis in the literature on media trust is questioned by some scholars who argue that declining trust in news media is closely associated with generally declining trust in institutions (Jakobsson & Stiernstedt, 2023). However, the instrumentalisation of a “lying press” by far-right political leaders has positioned trust in the news media at the centre of contemporary debates. Cross-country comparative research sheds light on the factors that seem to make people in some countries more resilient to mis- and disinformation (although much scholarship focuses on the United States and Europe). Distrust in legacy media has been associated with alternative news media consumption which is a cross-national factor associated with declining resilience to mis- and disinformation (Humprecht et al., 2023), and it is argued that high levels of trust in media can increase people’s resilience to mis- and disinformation (Humprecht et al., 2020). However, research also finds that trust in national news media does not build individual resilience when measured as a willingness to share, like, or comment on misinformation (in Canada, France, and the United States, but not the United Kingdom). In one

study, engagement with public service media (e.g., BBC, France Télévisions, or Canadian Broadcasting Corporation) was not found to build resilience in the short term (Boulianne et al., 2022). A comparison of 18 Western democracies using indices of populism, polarisation, media trust, shared media, strength of public service broadcasting, social media, and market size identified three groups of resilience to mis- and disinformation.



Resilience Cluster 1: High resilience to mis- and disinformation – Northern and Western European countries (Austria, Belgium, Denmark, Finland, Germany, Ireland, the Netherlands, Norway, Sweden, Switzerland, and the United Kingdom) and Canada; with countries seen as supportive of news media, more prone to political consensus, less polarised, and less prone to populist communication, and characterised by high levels of trust in media, shared media consumption, and strong public service media. Experience in the United Kingdom during and post Brexit when polarisation increased and mis- and disinformation flourished indicates risk in this “high resilient” cluster.

Resilience Cluster 2: Southern European countries (Greece, Italy, Portugal, and Spain) characterised by low resilience, with high levels of polarisation, populist communication, and social media news use, and low levels of trust and shared media consumption.

Resilience Cluster 3: The United States – a low-trust, politicised, and fragmented political and media environment (Humprecht et al., 2020). A follow-up study indicated that resilience was partly country-specific and highly dependent on the political and information environments (Humprecht et al., 2023). Focusing on Belgium, France, Germany, Switzerland, the United Kingdom, and the United States, chosen for their diversity in terms of resilience factors, resilience was measured as an inclination to ignore or disregard disinformation as opposed to engaging with it.



Apart from a limited number of cross-national indicators of resilience (i.e., heavy social media use, the use of alternative media, and populist party support), other variables, such as extreme ideology, populist support, age, level of education, and gender, vary by country. The contextual nature of resilience was validated in another study focused on awareness of, exposure to, and sharing of misinformation (Boulianne et al., 2022). The corrosive effects of mis- and disinformation on attitudes toward the news media may be less serious than is often assumed (Allen et al., 2020).

When trust is low or unevenly distributed over political divides, there are no easy solutions (Skovsgaard & Andersen, 2020). Journalists and news media organisations seek to remedy declining trust in news and industry trustworthiness in various ways (Kohring & Matthes, 2007). News media organisations address declining trust by attending to four aspects of news production (Kohring & Matthes, 2007; Prochazka & Schweiger, 2019) (see Figure 2.1). One is for journalists to be more transparent about their work, reporting factual content and avoiding opinions where possible (Newman & Fletcher, 2017). The benefits of sharing information on the process of writing a news story or sharing sources are illustrated by a cross-country initiative to increase transparency – the annual International Journalism Festival held in Perugia, Italy, the biggest free and open-to-the-public media event in Europe.

Figure 2.1 Building trust in news



SOURCE: Banerjee et al., 2023: 4; redesigned by Matthieu Mercier and adapted by Nordicom.

In Brazil, India, the United Kingdom, and the United States, based on survey research, qualitative interviews, focus groups, and other techniques (Banerjee et al., 2023), research suggests that news organisations should build trust and resilience by working towards more transparency, achieving better alignment of editorial coverage with concerns in people’s everyday lives, and preserving the media’s independence.

Conclusion

This chapter has emphasised the crucial role of news media in sustaining a robust public sphere, and attention focused on both legacy and online news media. We focused first on the increasing dependence of news media organisations on Big Tech–owned digital platforms, demonstrating that this varies by country, type of news, legacy versus online news, and the strategies news organisations adopt to address their audiences. We highlighted the financial precarity of many news media organisations in the face of Big Tech digital platform business models that prioritise audience engagement and operate in untransparent ways resulting in declining advertising revenues for news organisations. We also emphasised the need for independent news media and for supporting independent public service media (or public media) that aspires to contribute to reasoned debate in the public sphere.

The chapter examined research on trust in news media, showing that trust depends on multiple variables (e.g., age, gender, education, ideology and partisanship, and socioeconomic status) and that there are large differences across countries. We explained why the theories and research designs and methods used to study the relationships between news media engagement, attitudes, and behaviours, and the news media's role in the social and political ordering of societies, give rise to divergent perspectives on the impact of news media and mis- and disinformation.

This chapter also focused on changing journalism standards and practices in the news media industry and on research highlighting changes in media use, why people avoid the news and why some are more resilient to mis- and disinformation than others, and on the actors that engage in weaponising information in the public sphere. The various claims about the impact of mis- and disinformation on political polarisation were examined, as were measures designed to strengthen trust and resilience to mis- and disinformation. Overall, the chapter presented a complicated picture of what happens when trust in news media declines and people cannot tell the difference between accurate and inaccurate – or false – information. Lastly, we emphasised that much of the evidence on the role of news media in democracy is biased towards experience in the Global North and a Western understanding of news media's normative goals.

Artificial intelligence, information ecosystems, and democracy

Introduction

Humans are social beings. They communicate to achieve common goals based on convictions they develop through the information they receive and share. Democratic decision-making processes cannot function without public discussion of questions of general interest and sharing of ideas and debate about proposed courses of action and past decisions. These processes have become heavily digitalised (taking place in online spaces) and mediated (taking place in, and under, the rules, practices, and algorithmic systems of privately owned communication spaces). These spaces – information ecosystems – have rules, just as offline spaces do. In offline public spaces, laws set by states and enforced by executive power define the rules for public debate.

In online settings, the rules under which communication takes place are set primarily by private actors, such as the Big Tech owners of digital platforms, within the limits of what the laws allow. These actors enforce their communication rules through systems for content moderation that determine if the content adheres to the rules and curation that establish how the content is directed towards the platforms' users. The more platforms seek to automate these systems using artificial intelligence (AI), the more they impact online communication processes and, ultimately, influence democratic discourses and democracy. The integrity of information ecosystems therefore depends on an environment that favours transparency and accountability (Nowotny, 2021; Puddephatt, 2021).

In examining how AI systems (including Generative AI, or GenAI) and content governance (including content generation and moderation and curation) become interdependent, in this chapter, we assess the impact on information integrity as well as the broader implications for societal resilience and for social and environmental sustainability. We begin by clarifying what is meant by “AI” and how various uses of large language models (LLMs) impact fundamental rights, including freedom of expression and information,

privacy, and democratic participation. We then turn to assess developments in content governance, focusing on tools and processes that employ AI systems to generate content and to moderate and curate the content that is distributed online. We next address selected human rights issues that arise in the context of automated content governance and that have an impact on democratic decision-making processes (the details of legislation are taken up in Chapter 5). The analysis in this chapter confirms that an understanding of AI systems is essential if flows of mis- and disinformation are not to become a greater threat to the integrity of information within information ecosystems.

Defining AI

No single definition of AI is accepted in the policy and academic literature (Samoili et al., 2020). The European Union’s Artificial Intelligence Act of 2024 (AI Act), for example, defines AI as,

a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. (EC, 2024c: Art. 3.1)

The OECD has adopted a similar definition of AI as,

a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment. (OECD, 2022c: 7)

Neither of these definitions claims that AI systems emulate human intelligence. Instead, the focus is on functional capabilities that derive from using machine learning algorithms that work by identifying patterns in data. This interpretation is reinforced in a European Commission Joint Research Centre report:

AI is a generic term that refers to any machine or algorithm that is capable of observing its environment, learning, and based on the knowledge and experience gained, taking intelligent action or proposing decisions. There are many different technologies that fall under this broad AI definition. At the moment, ML [machine learning] techniques are the most widely used. (Annoni et al., 2018: 18)

This definition makes explicit the technologies – that is, algorithms and machine learning – that constitute AI systems, which other definitions gloss over. It is a reminder of the need to “look under the hood”, to challenge “the thingness of AI and its status as a stable and agential entity [...] To let

the term pass is to miss the opportunity to trace its sources of power and to demystify its referents” (Suchman, 2023: 1).

Pasquinelli’s (2023) *Social History of Artificial Intelligence* emphasises that these technological systems are not “intelligent” and discusses how research has been motivated by interests in labour-saving and surveillance. The problems raised by anthropomorphising AI systems are widely discussed in the literature (Floridi & Nobre, 2024), and proposals for defining AI vary across fields of study. In the field of political communication, for example, AI has been defined as “the tangible real-world capability of non-human machines or artificial entities to perform, task solve, communicate, interact, and act logically as it occurs with biological humans” (Gil de Zúñiga et al., 2023: 2). Activists and critical scholars emphasise the importance of focusing not just on technology, but also on politics, power structures, cultural narratives, and public perceptions (PublicSpaces International, 2024; Verdegem, 2021). Frau-Meigs (2024b) has argued for a new glossary of terms, such as “systems for statistical propositions”, to describe LLMs in the interest of encouraging more transparent discussion of the benefits and harms of technological advances. Mueller (2025: 2) has argued that the AI label is unhelpful because “AI” is part of a digital ecosystem that includes “computing devices, networks, data, and software”. Unless this is recognised, he has said, discussions about how AI should be governed are not meaningful.

There is not *an* AI; rather, there are different machine learning technologies, instances of which may be involved in processes related to information retrieval, synthesis, presentation, and governance. Machine learning technologies vary widely, ranging from relatively simple algorithms executing tasks (such as filtering for specific words) to deep learning algorithms (trained to assess the likelihood of content having been authored by an inauthentic actor, such as a disinformation bot) (Veale et al., 2023). While it is important to engage in critical discussion about what AI is, it is difficult to avoid using the term. This is because the proprietary nature of many of these technological systems means that details of the technologies are often not disclosed, and because this term is used widely in both research literature and popular and official discourses. In this book, we refer to “AI systems” and to the specific components of AI systems whenever possible.

AI systems are said to be “revolutionizing the way we create information across various mediums, including text, audio, images and video, presenting both challenges and opportunities” (Forum on Information and Democracy, 2024a: 18), and democratic control of these systems requires effective accountability structures for the whole AI systems lifecycle. Embedding AI systems within information ecosystems has an impact on content production (synthesising text, images, speech, and video), moderation (deciding whether content violates regulations), and consumption (deciding on the content’s audience). Concerns are raised about the potential for these systems to shape public discourse and, moreover, to do so in ways that have significant

implications for societal cohesion and resilience (De Gregorio & Stremlau, 2023). Thus, questions about the use of AI systems cannot be settled on technical criteria alone. A much broader range of issues must be addressed, including legal (e.g., does their use discriminate against certain groups?) and societal (e.g., does their use reduce the diversity of information available to publics?) questions (Katzenbach, 2021).

As the capabilities of AI systems advance and find new applications, it is anticipated “that algorithmic moderation and regulation will become more and more seamlessly integrated into our social lives” (Katzenbach, 2021: 6). As this process progresses, the increased “consumption and commodification of artificial intelligence applications in daily life”, coupled with the “extensive trust and reliance on these technologies in public and private sectors”, makes it essential to confront important rule-of-law questions (De Gregorio, 2023: 1). How these new technologies become embedded within people’s everyday lives is neither inevitable nor identical in different countries and regions. It is shaped and influenced by normative choices based on ethical values and societal (and international) goals that are pursued (or not), and by the experiences and outlooks of people in different regions and countries.

So-called GenAI refers to a broad category of machine learning systems that are capable of synthesising content (Bontcheva et al., 2024). They are typically trained on very large datasets and can generate content – synthetic media – in the form of text, images, speech, and video that is often difficult to distinguish in terms of quality from human-generated content (Martin & Newell, 2024). A notable example of the role of AI systems is OpenAI’s Sora, one of several GenAI systems capable of producing video footage from minimal text input with applications in the film industry, education, and marketing (Liu et al., 2024). LLMs first achieved public attention in November 2022 with the announcement of ChatGPT by OpenAI, and their use has significant implications for the public’s experiences of information ecosystems and the content that diffuses through them. LLMs play a role in creating realistic deepfakes and other types of mis- and disinformation. Despite being classified as GenAI, LLMs are simply statistical models of language use. While systems that use these models, such as chatbots, can produce plausible responses to queries, this should not be mistaken for natural language understanding. Thus, an LLM can be understood as “a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning: a stochastic parrot” (Bender et al., 2021: 617). The accountability of these systems is crucial, especially as AI-system developer companies, including OpenAI, are charged with putting “shiny products” on the market above safety, as is claimed by researchers who have left the company (Milmo, 2024). For the OECD, accountability can be defined in various ways. In developing metrics for trustworthy AI, the OECD defines

it as involving setting objectives and the functional specification, building an AI model to meet the specification, verification and validation, and its deployment, operation, and monitoring (OECD, 2023).

Some parts of the world are excluded from experiencing the benefits of these AI systems – “the fact that 2.6 billion [unconnected] people are not part of the digital world means they are not part of the AI world” (Bogdan-Martin, Secretary General, International Telecommunication Union, quoted in Vanoli, 2024: para. 32). Exclusion from the Internet – experienced disproportionately in the Global Majority World – keeps this population from using AI tools, including GenAI (Fendji, 2024). But these populations are not unaffected by AI systems, since they are recipients of information that circulates by other means. While the Global North deals with the effects of the growing pace of technological change, the Global Majority World struggles with access, inequalities in investment in online safety and content governance resources, poor infrastructure, and weak technology literacy levels (De Gregorio & Stremlau, 2023). At the same time, there is increasing use of AI systems in the Global Majority World, accompanied by a lack of investment in the underlying infrastructure and content moderation capacities for smaller language communities and non-Global North cultures. This is coupled with “workforce extractivism” – the use of low-wage ghost workers for training AI models (Monasterio Astobiza et al., 2022) – which complicates efforts to ensure the protection of human rights as AI systems become more widely available.

AI systems and human rights

The rapid development of AI systems makes it all the more important to understand how human rights apply in the digital age. Problems are created for fairness as a result of algorithmic biases, for freedom of expression and information, as well as for privacy protection when the impacts of AI systems developments are considered, including the impact on democratic participatory rights when AI systems are deployed to manipulate information. Human dignity serves as the cornerstone of human rights. More than 30 years ago, the guiding principles of the Vienna Declaration on human rights emphasised the indivisibility, universality, interrelatedness, and mutually dependent and reinforcing nature of all human rights (OHCHR, 1993). Predating this, the Universal Declaration of Human Rights (UDHR) of 1948 committed states to the “promotion of universal respect for and observance of human rights and fundamental freedoms”, declaring these rights a “common standard of achievement for all peoples and nations” (UN, 1948: preamble) (see also Chapter 1).

Human rights are fully applicable in the age of digital transformations, although much work is needed to uphold them in practice. Alongside technology

innovation, there is no need for *new* human rights. As then United Nations high commissioner for human rights, Michelle Bachelet (2019: para. 34), concluded, technological change does not necessitate new human rights conventions, but rather, “adapting the way we use institutions and processes. [...] We can protect rights effectively only if we constantly fine-tune our processes to find the right mix of interventions”. All societal actors have human rights obligations, albeit to differing degrees. Private online communication platforms have duties under the so-called Ruggie Principles: the *Guiding Principles on Business and Human Rights: Implementing the United Nations “Protect, Respect and Remedy” Framework* (Ruggie, 2011). Private entities must protect, respect, and provide remedies for violations of human rights under the overall control of states. Following international human rights law, states are expected to protect, respect, and ensure these rights for anyone within their control or jurisdiction (Fischer-Lescano, 2016), although, in the absence of regulation, these duties are not necessarily binding for all actors.

While Big Tech companies tend to frame their missions in human rights terms, such as giving people a voice on their digital platforms or protecting expression, research focuses primarily on potential human rights violations by governments and less on areas where platform business models might negatively impact user rights (Jørgensen, 2017; Kettemann & Schulz, 2023). In the light of a reluctance to commit to substantial transparency obligations over the decades, regional and national approaches have emerged to apply human rights obligations more directly to platforms (Müller & Kettemann, 2024). The first big-picture approach can be seen in the European Union’s *Digital Strategy*, which attempts to curtail the influence of large digital companies by imposing obligations on them that mitigate the negative effects of online communication and, at the same time, to promote the implementation of fundamental rights (EC, 2022b). The European Union has emerged as a key international norm-maker for the digital arena, sometimes referred to as the “Brussels Effect” (Bradford, 2020). Legislation, including the AI Act, provides some substantive obligations, but through stringent transparency and compliance obligations (EC, 2024c; Müller & Kettemann, 2024; Werthner et al., 2024).

United Nations initiatives, such as a March 2024 General Assembly Resolution, show there is awareness of technology’s role in both contributing to disruptive change and having the potential to build bridges within and between countries. The Resolution emphasises that trustworthy AI systems for sustainable development should be promoted globally in line with existing human rights obligations (UN, 2024c, adopted by the UN General Assembly 21 March 2024). By September 2024, AI systems had been positioned with other frontier technologies as a means to “turbocharge development”, securing a place as Objective 5 of the United Nations’ *Global Digital Compact*, which emphasises the need for a “balanced, inclusive and risk-based approach to the governance of artificial intelligence (AI)” (UN, 2024b: 50).

Algorithmic bias and fairness

Algorithmic bias involves systematic errors within AI systems that lead to unfair results (Hasimi & Poniszewska-Maraña, 2024). Unfairness can be understood as privileging members of one group over another without adequate reasons. The biases of personalisation systems and search engines have been recognised and demonstrated empirically for at least a decade (Eubanks, 2018; Rieder & Sire, 2014; Rieder & Skop, 2021). When these systems are used in settings where automated decisions impact individual or collective rights or values, these biases can lead to unfair and untransparent outcomes, not least because of economic incentives to favour results consistent with corporate interests. Algorithmic decision-making in areas such as employment, law enforcement, and lending can disproportionately negatively affect marginalised communities and contribute to their exclusion from participation in democratic processes or the full enjoyment of their rights (Baecker et al., 2023). Research predominantly in the Global North, but also in the Global Majority World, reveals how algorithmic bias can lead to decisions by law enforcement authorities that disproportionately penalise minority ethnic groups and immigrant communities (Chouliaraki & Georgiou, 2022; Gurumurthy & Chami, 2019).

Biases arise from various factors linked to how, by whom, and in which institutional or organisational setting an AI system is developed, particularly regarding the data used to train it – for example, when training data is incomplete or contains historical prejudices or assumptions that are then replicated: if, in the text on which an LLM is trained, doctors are primarily described as male, then answers generated by the LLM will replicate this (Belenguer, 2022). Similar replication of stereotypes has been shown to happen in image-generating LLMs. Even AI systems trained using what is believed to be unbiased data may produce biased outputs, since a lack of transparency in how their outputs are produced makes it difficult to exercise effective oversight over their performance (Pollicino & De Gregorio, 2022). The personalisation algorithms used on social media platforms to decide what content users are exposed to exploit the data that users create when they interact with content. Once human activity is datafied in this way, AI can be used to model user behaviour, and the model can then be applied in ways that are biased towards the interests of platforms, leading to the promotion of content that maximises user engagement at the expense of quality and veracity (Pfeiffer et al., 2023). These are all consequences of the way LLMs use their training data to synthesise outputs based on statistical prevalence – reducing the diversity of inputs into the specificity of a single output. In addition, LLMs may be trained on synthetic data, that is, “data that mimic and substitute empirical observations without directly corresponding to real-world phenomena” (Offenhuber, 2024: 1; see also Bonfanti, 2020). Critical assessments of the use of such data in training AI

models may be helpful in protecting privacy and improving datasets that have a representational link to the “real-world”, for example, addressing biases, but the new generation of AI models is not yet capable of explaining why a given output has been generated.

Algorithmic fairness refers to the aspiration of creating and implementing AI systems that do not discriminate or bias against specific persons or groups based on protected characteristics, such as race, gender, or ethnicity (Ferrara, 2024a; Johnson, 2023). Fair AI algorithms would make decisions without favouring one individual or group over another (Hall & Ellis, 2023). To achieve this, attempts are now being made to increase the quality of training datasets. One approach is based on the premise that greater diversity in training data will reduce the prevalence of bias. As an example, IBM launched a Diversity in Faces dataset to help overcome specific biases in facial recognition technology (Smith, 2019). This dataset includes a million images of faces annotated with details that provide a broad representation of human faces, such as age, gender, skin tone, and facial features drawn from many different countries and cultures. By using this dataset, developers can train facial recognition systems that are less likely to reproduce stereotypes regarding certain groups (Wiggers, 2019).

Diversity in training data is expected to contribute to mitigating the risks of bias in AI systems that use these models. Diversity in development teams can offer a variety of perspectives that challenge conventional norms and biases that may be overlooked in more homogenous teams. The setup of development teams – and those working on AI ethics generally – is substantially linked to product design decisions (Martin, 2022). Microsoft has embraced this strategy through its Inclusive Design Initiative, which employs people with diverse backgrounds (including disabilities) to design and test new products (Microsoft, 2023). Evidence of the effectiveness of such corporate diversity strategies is inconclusive, and, in some cases, no direct association has been found between the sociodemographic diversity of AI systems developers and AI systems output biases. The viewpoint diversity of those holding machine learning, coding, or data analyst jobs is found to play a much stronger role, based on a relatively small-scale study (Chi et al., 2021; Harris, 2023; Park, 2024). Various forms of discrimination are likely to persist in the prevailing culture. This is likely to be encouraged if leadership is skewed to favour certain groups, as illustrated by high-profile resignations from some of the Big Tech companies.

In many countries of the Global Majority World, AI systems development and deployment are at a nascent stage, potentially allowing these countries to design robust anti-discrimination rules before broad uptake (Gurumurthy & Cham, 2019). For the Global Majority World, questions about “human rights, democracy and autonomy in the countries of the majority world are not trivial” (Ricaurte, 2022: 732). For example, the development of fair AI systems may be hindered by the limited availability of training data in many

Global Majority World languages (Ricaurte, 2022, citing Horowitz, 2021; HRW, 2023), and Timcke and colleagues (2024) have shown that Africans are not represented in AI training data. The challenges around guaranteeing fairness will also increase as AI systems progressively become enmeshed in the processes that define the social conditions in which meaning is produced. These, in turn, are dependent on the level of trust in them, their prevalence, and institutional roles (Pfeiffer et al., 2023).

Freedom of expression and information

Freedom of expression is a cornerstone for the formation of democratic societies and, as such, is protected by all human rights instruments (UN, 1948, 1966) and all regional human rights conventions. This right includes the freedom to express and hold one's own opinions, to impart information, to seek and receive information, and, implicitly, freedom of media expression. Given the technological realities of online communication, the right to freedom of expression is implicated in other rights, such as the right to health (seeking and imparting health-related information) and to education (seeking and imparting information related to education, attending classes, and accessing research papers). AI systems allow for much easier access to online communication spaces and information interchange, and they have an impact on what information can be seen (Dias Oliva, 2020). All platforms use AI systems to govern online communication and optimise user engagement (Gillespie, 2020; Longo et al., 2024), and there is a substantial impact of these content governance systems on freedom of expression across information ecosystems (De Gregorio & Dunn, 2023).

Information enables individuals to make educated judgments by helping them become acquainted with facts and societal issues (Masur, 2020). It is a crucial component of individual liberty. Nevertheless, people's ability to obtain and interpret information can be restricted when they encounter mis- or disinformation or biased content. Another challenge people face is a lack of reliable, accurate information, a problem sometimes compounded by information overload, and worsened when there is a decline in global trust in news, as we emphasised in Chapter 2 (Samoilenko & Suvorova, 2023). Focusing on tweaking content governance practices and systems ignores the underlying causes of social discord and distrust that give rise to polarised public opinion. Some argue that focusing on the public worthiness of information, rather than on information "disorder", can better reveal the complex elements of visibility, access, reflexivity, mediation, influence, and information legitimacy. Improved insight into how these elements combine in different ways to foster new imaginings of publicness are important if democracy is to flourish (Geiß et al., 2021; Splichal, 2022a).

Privacy protection

AI systems also present significant challenges to people's rights to data protection and privacy. The comprehensive collection and analysis of data by these systems leads to the development of multidata, points-based profiles of individuals, often without their explicit consent, which can separately and in aggregate violate their right to privacy (Bontridder & Pouillet, 2021). For example, Meta has said it is extending the jurisdictions in which it collects public data to train its models beyond the United States, although Data Protection Authorities in the European Union and Brazil and beyond have sought to stop this practice. In other countries, such as the United Kingdom, after some changes, this practice has been deemed a legitimate interest in processing data (Forum on Information and Democracy, 2024b). Moreover, the way consent is obtained for data collection does not often meet the threshold of being informed. Many users experience consent fatigue, agreeing to privacy policies without understanding the implications (Abdulrauf & Dube, 2024; Avle, 2022; Barocas & Nissenbaum, 2014; Richards & Hartzog, 2019; Turow et al., 2023).

As most online communication takes place in private communication spaces that are financed through data collection, there is an incentive for platforms that use automated content governance tools to configure them in a way that maximises data collection. This can lead to interferences with, and violations of, rights to privacy and data protection. These can be addressed to some extent by enforcing existing privacy and data-protection laws and international human rights standards that emphasise consent, data minimisation, and purpose limitation in data processing – for example, Article 12 of the UDHR and Article 17 of the ICCPR protect privacy and personal data (UN, 1948, 1966).

Training an LLM on personal data effectively encodes aspects of this data into the model's parameters. This depends on whether training involves access to personal data and where this data is stored. In the case of ChatGPT, additional data is uploaded to OpenAI's servers, and the company retains this data, but some LLMs allow for data to remain under local control. Even where there is no direct retention of the data, the model learns patterns and information during training that may be reconstructed or inferred by analysing its output. Studies have shown that it is possible to extract specific data points from LLMs through techniques like model inversion or membership inference attacks, where queries to the model can reveal whether certain data was used in training (Jagannatha et al., 2021). However, this approach may be limited by the fact that it suggests that it is possible to invert outputs to inputs, ignoring that the model combines inputs according to probabilistic weights that are derived from a combination of inputs, rather than linearly from any single input. In addition, models learn and change when model–user interaction or “user embedding” occurs in addition to learning in response to user text prompts (Ning et al., 2024).

There are several potential technical solutions and strategic reforms that can be implemented to address the privacy risks posed by LLMs and other types of AI systems. These aim to enhance privacy protection, ensure transparency, and uphold ethical standards within AI systems development and deployment (Lepri et al., 2018; Ong et al., 2024; Yan et al., 2024). Regarding technical solutions, protection from de-anonymisation risks can be achieved by using differential privacy methods that add random noise to the data in a way that prevents the identification of any individual in the dataset, and these methods have been adopted by companies such as Apple and Google (Zhao & Chen, 2022). However, surveillance, defined as the “process of observing individuals or groups for a purpose and mak[ing] inferences/judgements on their behavior” (Fontes et al., 2022: 2) is transformed by the scope and scale of AI-enabled datafication, that is, the quantification of people’s everyday activities in real time by digital platforms – sometimes known as dataveillance (van Dijck, 2014). AI systems are used to identify patterns of behaviour in data, leading to new risks and potential remedies, including pressures on AI developers to adhere to ethical standards.

Establishing and adhering to ethical standards when developing AI systems is essential to mitigate risks related to privacy, bias, and other potential harms. For instance, the Institute of Electrical and Electronics Engineers has proposed an ethical framework for AI and autonomous systems that includes guidelines for prioritising human well-being, data agency, and accountability in AI systems (IEEE, 2019; Gunkel, 2024). Similarly, the Partnership on Artificial Intelligence to Benefit People and Society (Partnership on AI), which includes stakeholders from various organisations, promotes best practice in AI development, focusing on fairness, transparency, and accountability to ensure AI systems are used responsibly, although substantial changes in company policies or product priorities have not materialised (Borocas et al., 2023; Caton & Haas, 2020). UNESCO (2022c) has also published a recommendation on the ethics of AI.

Democracy and participatory rights

While United Nations member states are committed to democratic governance, most of the world’s population lives in states that suffer from democratic deficits. A 2023 analysis of the state of democracy globally concluded that it is “complex, fluid and unequal” (International IDEA, 2023: Executive summary, p. 8). The Vienna Declaration and Programme of Action of 1993 clearly links democracy with human rights, urging member states to continuously foster democratic principles to enhance human rights protection (OHCHR, 1993). Democracy fundamentally relies on the principles of free, equal, secret, and independent elections and democratic decision-making processes. Particularly problematic are targeted mis- and disinformation campaigns that aim to manipulate elections and stir conflict by weaponising information,

as discussed in Chapter 2. The manipulation of democratic decision-making processes is facilitated by AI systems, and even non-subliminal techniques can be manipulative, especially for vulnerable groups, such as children, if they exploit mental health vulnerabilities, immaturity, or lack of digital literacy. As AI systems evolve, the potential for misuse increases.

The utilisation of AI systems in political campaigns and election processes leads to concerns about the transparency, accountability, and manipulation of democratic decision-making. The electoral landscape faces significant risks from the very tools that enable campaigns to target voters with exceptional accuracy: the capacity to disseminate false information, to manipulate perceptions through microtargeting, and to magnify controversial content (Schippers, 2020). A comprehensive Norwegian review of the role of AI systems, social media, and cyber operations in elections held in 2024 found that “there is no available overview of the reach and potential of such technology, nor of how it will impact on society and individuals over time” (Ministry of Local Government, 2025: 7).

The capacity of AI chatbots, for example, Microsoft’s Bing Chat, was tested over several months during elections in Germany and Switzerland. This GenAI chatbot produced factual errors to queries on election topics with a near 30 per cent error rate (Helming, 2023). The use of AI systems to personalise content on social media platforms has the potential to sway voters and create divisions in public opinion, affecting individuals’ abilities to freely engage in government and public affairs. With advanced data analytic capabilities making voter microtargeting significantly more accessible, this has the capacity to enhance engagement and voting percentages, but it also exposes voters to manipulation via hyper-targeted content that can seek to sway their opinions or even discourage them from voting (Michael, 2023). The next section examines approaches to content governance which are evolving alongside these developments in AI systems.

AI systems and content governance

AI systems that are deployed by Big Tech companies on their digital platforms manage the visibility and spread of information, including mis- and disinformation (Sančanin & Penjišević, 2022). This section addresses content generation and governance – that is, content moderation, distribution, and amplification. It also assesses the impact of AI systems on information ecosystems and discusses how AI systems are being used in the news media industry and by mis- and disinformation actors.

Social media platforms have become vital arenas for public debate, where users gather information, share ideas, and form opinions. Content governance systems impact these processes because they frame the conditions under which content is seen and with whom it is shared (Jungherr & Schroeder, 2023). These systems utilise user behaviour, previous choices (interest histories), and

past interactions to customise content streams, control content visibility, and enhance engagement metrics. One objective of AI-based content governance systems is to reduce the prevalence of undesired content such as mis- or disinformation, including hate speech and propaganda (Christodoulou & Iordanou, 2021). Importantly, their design, implementation, and accountability lie in the hands of the platforms where they are used. These systems, and the governance policies and practices they are intended to support, vary from platform to platform and from jurisdiction to jurisdiction, changing over time, especially with changes in ownership, as illustrated in the case of X/ Twitter (Burkart & Huber, 2021; see also EC, 2024d for a demand that X explain its content moderation compliance with European Union regulations).



Twitter's transformation under Musk

On his takeover of Twitter in late 2022, Elon Musk announced that the reason he acquired Twitter is because it is “important to the future of civilization to have a common digital town square, where a wide range of beliefs can be debated in a healthy manner, without resorting to violence” (York, 2022: para. 13). Not long after, he introduced significant changes to Twitter’s content policies and practices, signalled by reinstating some high-profile users who had been banned for violating the platform’s misinformation and hateful conduct policies (Ivanova, 2022). Changes in practices were inevitable with the dismissal of a large proportion of staff responsible for human rights, AI ethics, trust, and safety (Brewster, 2024; eSafety Commissioner, 2024a). X introduced Community Notes, which aimed “to create a better informed world by empowering people on X to collaboratively add context to potentially misleading posts” (X, 2024: para. 1), but X retains control over which contributions are approved and made visible to users. The Australian eSafety Commissioner (2024b: para. 9) has criticised X for letting the worst offenders back online, “while at the same time significantly reducing trust and safety personnel whose job it is to protect users from harm”.



The AI content governance systems that drive social media platforms are designed to enhance user engagement by personalising online experiences, and they are, in principle, neutral on the veracity of content. As a consequence, if mis- or disinformation content provides the most engagement, then the system – if not properly reviewed – will increase the dissemination of such content (Bontridder & Pouillet, 2021; Ohme et al., 2024; Reisach, 2021).

AI systems in content generation

The availability and ease of use of GenAI has arguably “democratised” content production. Without specific detailed technical know-how, users can now create digital audio, video, or text content cheaply, using a wide range of apps, and distribute this content through digital platforms (Allen & Weyl, 2024; Cooke, 2023). With this comes potential side effects, which stem from the increase in volume, velocity, and potential persuasiveness of problematic content and its decreasing cost of production (Feuerriegel et al., 2023).

The owners of digital platforms have started to address the challenges of text, images, speech, and video produced by GenAI, but, in jurisdictions without rules for risk assessment obligations, the internal rules are often vague or inconsistently enforced: “The driving force is either the misleading and harmful potential or a more compliance-oriented approach in terms of copyright and quality standards of the content” (Miguel & Krack, 2023: 3). Given that audiences find it difficult to distinguish between GenAI- and human-produced content (Kreps et al., 2022), it is important to raise levels of AI literacy (see Chapter 5) and to impose disclosure obligations for AI-generated content or AI-operated accounts. So far, however, despite high-profile cases, there is no evidence that GenAI is *systematically* used as a tool to synthesise politically motivated mis- and disinformation (Kreps et al., 2022; Simon et al., 2023). Even if this is the case, there is no doubt that it is being used with growing pressures on tech companies, prompting them to sign a voluntary accord in early 2024 to prevent AI systems from disrupting elections (O’Brien & Swenson, 2024).

Copyright is a challenging issue for AI-generated content because it applies both to the data used for training and to the generated output. Training AI systems, especially LLMs, often involves ingesting vast amounts of text harvested from the Internet, much of which is copyrighted (although some exceptions exist), raising questions about whether this usage constitutes “fair use” or exceptions to copyright depending on the jurisdiction, or requires explicit permission from rights holders (Kretschmer et al., 2025; Miltner & Highfield, 2024; US Register of Copyrights, 2025). The output generated by AI systems also sometimes (and inexplicably) reproduces copyrighted material verbatim (US Register of Copyrights, 2024). In addition, there are concerns about the privacy implications of the scraping of vast quantities of personal information by AI developer companies (Solove & Hartzog, 2024).

AI systems in content moderation and curation

AI systems are increasingly used by platforms for implementing content governance guidelines on how content is sourced (or created) and then distributed. Content moderation involves identifying and removing or flagging inappropriate, harmful, or illegal content based on predefined criteria. This

definition of criteria – the setting up of internal standards and community guidelines – is a powerful act, which, coupled with algorithmic content moderation and curation (i.e., governance), gives the Big Tech owners of digital platforms a role that researchers call the “arbiters of truth” (Schaaake & Fukuyama, 2023). It is not so much truth that is decided on, however, but what content stays on a platform and what content is given more visibility. Content curation systems then select and organise content that has passed the moderation stage for distribution. These systems are used by platforms to determine who sees what content, often personalising it by matching against users’ preferences, as revealed by modelling their past behaviours (Gillespie, 2020). The use of these systems takes place within the framework of existing and new laws shaping platform behaviour, including rules for transparency and user rights. No moderation or content curation system is neutral or non-discriminatory. If it did not treat content differently, it would not be doing its job. Certain categories and procedures must be used to structure the content presented to social media users. Choices must be made, even if the choice is to present content in chronological order. As a report for UNESCO’s regional office in Montevideo put it:

AI technologies are not neutral; they inherently reflect the values of their developers and the broader development and deployment ecosystem. While they have the potential to enhance accountability in public institutions and their representatives, foster greater participation and pluralism to enrich citizen engagement, and make democracy more inclusive and responsive, they can also amplify autocratic tendencies and be used for potentially malicious and manipulative purposes. (Innerarity, 2024: 10)

It is essential to puncture the “‘fallacy of AI neutrality’ – represented by the mistaken belief that AI systems can be designed in an inherently unbiased and neutral manner” (Verhulst, 2023: 1). Research shows that content moderation and curation systems suffer from biases and encode non-transparent decision-making processes. They are optimised for engagement, that is, to personalise and distribute content to audiences that the system predicts will engage them meaningfully (Sančanić & Penjišević, 2022). This means that they can be designed and deployed to achieve political ends and in ways that exacerbate individual and societal risks. GenAI content is distributed in this way, even though a substantial number of platforms do not have sufficiently detailed policies in place – and may not adhere to them when they do.

There are concerns about the political sensitivity of LLMs and their potential to deepen societal divisions (Biju & Gayathri, 2023). These tools are relatively new and are updated quickly, but research shows that the output of three LLM-based chatbots (ChatGPT, Bing Chat, and Bard) exhibits varying degrees of bias in response to political queries concerning authoritarian regimes. This is influenced by the language of the prompt. Significant disparities have been found regarding chatbot answers, with Russian-language queries resulting in

evasive answers regarding content that can be viewed as critical of Russian authorities. Anecdotal evidence shows that this applies to similar queries in Mandarin Chinese on issues such as the persecution of Uyghurs (Urman & Makhortykh, 2025).

Human moderators still play a role in overseeing certain automated decisions and, depending on the jurisdiction in which a social media company operates, become active once a user requests that a content-related decision is reviewed. These cognitive assemblages involved in content moderation have been described as a “cobbled space of pre-emptive calculation” (Crosset & Dupont, 2022: 10). Generally, the trend clearly goes towards more automated moderation, especially in areas where the law is regarded as being clear, such as terrorist content (Haas & Kettemann, 2024; Macdonald et al., 2019). Where the law is less clear, as in the case of some forms of mis- and disinformation, automated tools focus less on content and more on markers related to the distribution channel or the behaviour of the account from which the content was launched (Bontridder & Pouillet, 2021). Researchers have criticised that even well-optimised moderation and curation systems can “exacerbate, rather than relieve, many existing problems with content policy” because they increase opacity, complicate “issues of fairness and justice in large-scale sociotechnical systems and [...] re-obscure the fundamentally political nature of speech decisions being executed at scale” (Gorwa et al., 2020: 1).

Since content produced by AI systems can exhibit and reinforce biases against historically marginalised and minority groups (Ross Arguedas & Simon, 2023), safeguards must be implemented to prevent these systems from intensifying existing societal inequalities, along with efforts to use these systems to help elevate the representation of underrepresented groups in the content produced (Forum on Information and Democracy, 2024a). Efforts to promote ethical standards and diversity in development teams are part of the solution but are not themselves sufficient. Gillespie and colleagues. (2023: 4) have called for an expanded research agenda on content moderation, arguing for grasping “the breadth and depth of moderation, across the entire ecosystem of content provision and deep into the infrastructural stack of distribution”.

The hyper-personalisation of content curation systems attracts much criticism. Some researchers fear that they may lead especially vulnerable media consumers, such as children and young adults, into rabbit holes of potentially harmful content, among many other harms (Amnesty International & AI Forensics, 2023). “Harmful” is a difficult criterion to use as a basis for assessing platform content policies. For example, there are few globally accepted examples of prohibited speech; much hate speech, for instance, falls under the protection of free speech in rules in some jurisdictions, such as the United States, and depending on the jurisdiction, there are different definitions of illegal speech (Gillespie, 2020; Galli et al., 2023). This is why it is sometimes

argued that automated systems would work better if there was global consensus or a largely agreed upon definition of what to find or filter, as in certain cases of terrorism and terrorism financing (Haas & Kettemann, 2024). Any such effort to forge consensus is likely to be disputed due to cultural and political differences and, even if achieved in the framework of international human rights obligations, may not be translated consistently into practice.

AI systems have also been used, for example, to improve crisis communication (Coombs & Holladay, 2022; Jin & Austin, 2022). A study by the Organization of Security and Cooperation in Europe (OSCE) suggested that states should mandate platforms to undertake “crisis-sensitive human rights due diligence”, “crisis-sensitive human rights risks and impact assessments”, and emergency measures. Any platform action should “consider proportionality and reliability on AI tools and automated measures” (Haas & Kettemann, 2024: 9). Globally, crisis-sensitive human rights approaches by private actors have been urgently demanded (Fatafta, 2024). Similar obligations are outlined in Europe’s digital rules, such as the Digital Services Act of 2022, which contains obligations for platforms to conduct risk assessments on the impacts of their rules and moderation practices on values, including societal cohesion, public health, and democratic decision-making processes (EC, 2022c).

Given that platforms use AI systems for content governance, it is best practice (and legally required in certain jurisdictions, such as the European Union) that they should inform users. However, research shows that users tend to trust moderation decisions less when they know they are automated (Ozanne et al., 2022). This showcases the complexity of achieving the responsible visibility of automated content governance, and user trust is also conditioned by education background and the sociopolitical setting (Kim & Moon, 2021). One approach is to increase meaningful oversight, including external control over algorithmic systems (Nahmias & Perel, 2021). This intervention into the private communication realm by platforms, governed by terms of service and algorithmic systems, can be legitimised by reference to the increasing impact of these norms and practices on public values that need to be integrated into the systems.

Meta’s Oversight Board is one of the early efforts to make the governance of a commercial platform more inclusive of external input (Kettemann & Schulz, 2023). The impact on Meta itself tends to be judged as largely positive, if not very effective, and the board has been described as overseeing “one of the largest speech systems in history” (Douek, 2024: 373). However, the board has not had substantial cross-industry influence and has been unable to substantially change the speech governance priorities that Meta exhibits (Ang & Haristya, 2024; Douek, 2024; Gulati, 2023). Boards such as Meta’s examine specific cases of content moderation judgments, and broader forms of oversight aimed at increasing accountability are limited by researchers’ access to relevant data (Nahmias & Perel, 2021).

AI systems and news media

Content created by GenAI is widely expected to benefit news media diversity by contributing to the efficiency of content generation in specific contexts, and by supporting journalists in optimising the circulation of content contributions after publication, since articles can be published in cross-media formats without substantial additional costs. Research points to the potential of GenAI to “synthesize broadcast videos using news text during a news broadcast” with better results than manual generation (Wu et al., 2023: 13). However, the implementation of these systems requires time and investment, and gains in efficiency and productivity should not be assumed (Simon, 2024; Simon & Isaza-Ibarra, 2023). As the use of GenAI becomes widespread, this can alleviate the burden of relying on overworked newsrooms by automating certain, more mundane, reporting tasks. However, the challenges of news organisations’ use of these systems must be addressed if trust in news media output is to increase and these organisations are to adhere to ethical standards of data collection and the principle of universality, in contrast to promoting personalised news and other content (Horowitz et al., 2022; Ross Arguedas & Simon, 2023; Vaccari & Chadwick, 2020). AI systems also have a bearing on freedom of expression when they influence editorial decisions, especially when there is a conflict between the editorial need for autonomy and goals that AI tools are optimised for (Helberger et al., 2020).

Concerns have been expressed that the adoption of AI tools by news media organisations for content creation will lead to a growing dependency (Simon, 2022) (see Chapter 2), illustrated by AI guidelines for media in 17 countries (de Lima Santos et al., 2024). A survey published in 2023 indicated that GenAI tools, such as ChatGPT, were being used in 49 per cent of newsrooms worldwide (WAN-IFRA, 2023). Countries in the Global North and China are leading innovation in AI newsrooms, but research on adoption mainly focuses on the Global North (Beckett, 2019; Beckett & Yaseen, 2023; Kothari & Cruikshank, 2022; Marconi, 2020). Once content is created, news organisations become increasingly dependent on the AI systems used by digital platforms for distribution or circulation. This raises the need for attracting audience traffic that is stimulated by algorithmic personalisation. The effects of the interaction between audience traffic and the means to increase the flow of this traffic have implications for the production and visibility of content.

The question of who or what curates online content takes some of the power out of the hands of journalists, the traditional gatekeepers. Platform selection mechanisms usually involve a combination of algorithmic curation (based on criteria specified by business managers) and human editors, making it unclear what the core values underlying selection decisions are, and to what extent they reflect democratic principles (van Dijck et al., 2018b). The impact of curation systems is especially sensitive in public service media environments, where there is a mandate to reach a broad public (Horowitz et al., 2022). Research conducted in France, Germany, Greece, Italy, Poland,

and Sweden emphasises that the news media's growing dependence on algorithms means that those who access news media online to meet their information needs do so despite their concerns about the risk of encountering mis- and disinformation (Schaetz et al., 2023). Certain platforms have started to deprioritise news and favour more personal or emotionalising content (Meese & Hurcombe, 2021). As indicated in Chapter 2, weaker distribution of accurate information is associated in some studies with more polarised and polarising media consumption behaviour (Schirch, 2021). Platform algorithms using AI tools are playing an increasing role in shaping news distribution (Meese & Hurcombe, 2021; van Dijck & Poell, 2013). Some news organisations depend heavily on online traffic driven by third-party digital services, leading to dependency on social media for news distribution, a trend that does not uniformly affect the entire industry (Bakke & Barland, 2022).



Facebook's interest in news content has grown as it has sought to monetise online advertising and counter X/Twitter's emerging status as a key news source. In 2013, the company began promoting news publishers' content in its personalisation system. This encouraged news organisations to focus on Facebook distribution strategies for their news. Facebook developed technologies for hosting content directly (e.g., the launch of Instant Articles), and incentivised publishers to keep their content on its platform. The platform's shift to video content and the introduction of Facebook Live led the media industry to adapt to these changes. The relationship between publishers and Facebook soured due to monetisation challenges, inflated video metrics by Facebook, and controversies surrounding mis- and disinformation, especially during the 2016 United States presidential election. Facebook's response was to step away from news distribution in 2018, changing its News Feed algorithm to prioritise personal content. Faced with this challenge, some news media organisations altered their distribution strategies, aiming to regain control of revenue streams and favour core audience interests over Facebook demands (Meese & Hurcombe, 2021).



The extent to which the push to adopt AI tools will increase news media dependency on the affordances of digital platforms is unclear (Simon, 2022). Claims that the "AI goldrush" will increase the potential for infrastructure capture and shift even more control to platform companies raise questions about control, dependence, and autonomy as the adoption of AI tools in newsrooms extends platform control over the news production processes and

the distribution networks (Simon, 2022). While there is a growing market for AI tools to cater to newsrooms’ needs, with smaller players such as Narrativa, Retresco, Adobe, and others trying to position themselves in the market, the dominant players are operating in an oligopolistic market. Table 3.1 illustrates the growing prominence of AI systems use in the production and distribution processes of newsrooms.

Table 3.1 AI system use in news media gatekeeping

Production and distribution process	Use of AI systems
Access and observation	Information discovery
	Audience and trends analytics; story detection
	Prompting for new ideas following from a news story
Selection and filtering	Verification, claim matching, and similarity analysis (e.g., for fact-checking)
	Content and/or document categorisation; analysis of datasets
	Automated collection and analysis of structured data (e.g., financial, banking, and sports data)
	Coding assistance for various tasks
	Transcription and translation of audio and video
	Search in archives and/or metadata
Processing and editing	Brainstorming and ideation
	Content production (writing of draft text or articles; editing of news content)
	(Re-)formatting of content for online, social media, print, broadcast (e.g., summarisation, simplification, stylistic changes; text-to-video, speech-to-text, text-to-speech translation)
	Copy editing, adaptation to house style
	Tagging of content, headline, and SEO [search engine optimisation] suggestions
Publishing and distribution	Personalisation and recommendation
	Dynamic paywalls, audience analytics
	Content moderation

SOURCE: reproduced from Simon, 2024: 13

Multiple factors influence the extent of news organisations’ dependence on digital platforms and their AI tools (Poell et al., 2023; Pyo, 2022; van Dijck et al., 2018a). These include the country (e.g., there is weak evidence of dependence in Germany); the kind of news organisation; whether organisations are established, legacy, or digital-only (e.g., a study in South Korea found that legacy organisations experienced greater pressure than digital-only); and how public service media organisations address their public role and audience reach (e.g., to young people) (Hase et al., 2023; Eichler, 2023).

The wave of enthusiasm surrounding AI systems centres around their potential to transform the social roles of journalism, especially as these tools support the profession's core functions in a democracy (Beckett & Yaseen, 2023; Lin & Lewis, 2022; Moran & Shaikh, 2022). Analysis of media coverage of AI system use in journalism over a five-year period in the United Kingdom and the United States indicates that opinions are far from uniform (Moran & Shaikh, 2022). There is a tension between the industry (newsroom leaders and funders) advocating for the use of this technology long-term, and professionals (journalists) highlighting concerns, for example, about the impact of AI systems on accuracy, fairness, and transparency (Beckett & Yaseen, 2023). The use of AI systems in journalism is normatively evaluated in relation to stages of news work – information gathering, selection and production, distribution and consumption – with attention to dimensions of accuracy, accessibility, diversity, relevance, and timeliness (Lin & Lewis, 2022).

Discussions on how to develop AI tools responsibly need to be grounded in a normative perspective on the underlying values and principles, including the need to start with identification of values and principles with multiple stakeholders; the development of a forward-looking vision on the role of journalistic AI, grounded in a normative framework focused on editorial mission, fundamental rights, and the democratic role of the media; and understanding how journalists, editors, managers, developers, users, and other stakeholders can be empowered to become active agents in decision-making processes around the implementation of journalistic AI. The argument for a more inclusive decision-making process comes from the realisation that AI is not just a tool, but an integral component of the public communication infrastructure, whose design is of concern to all stakeholders. The challenge is “to design decision-making routines so that they become more accountable to the public, more inclusive and cognisant of diverse and underrepresented voices in society, and less dependent on a small number of major technology companies” (Helberger et al., 2022: 1621).

Ethical concerns underlying the adoption of AI tools by journalists include whether automated content is consistent with editorial criteria; personalisation that respects diversity and promotes a thriving public sphere; monitoring the quality of data to respond to bias; responsible safeguarding of user privacy; quality journalism with an emphasis on the human factor; funding of platforms and journalism independence; and AI systems to foster the values of journalism (Pocino, 2021). The Council of Europe's Steering Committee on Media and Information Society published *Guidelines for the Responsible Use of Artificial Intelligence in Journalism*, and there are numerous codes of practice to guide the use of these technologies (Council of Europe, 2023). The inclusion of AI tools in journalism routines could shift moral and editorial responsibility away from newsrooms, with consequences for public perceptions of news media bias (Calice et al., 2023; Moran & Shaikh, 2022). A study of professionals in newsrooms in 16 countries in

the Asia Pacific, Europe, Latin America, the Middle East, and North Africa (MENA), as well as North America and sub-Saharan Africa regions, found that ethical concerns were significant. More than 60 per cent of respondents were concerned about editorial quality, and many expressed a desire for AI system transparency and the implementation of ethical guidelines (Beckett & Yaseen, 2023). In addition, developing AI system de-biasing techniques has been found to be challenging for journalists, with the capacity to address this issue depending upon the data quality that is available to journalists in their work (Dierickx et al., 2023a, b).

Use of generative AI by mis- and disinformation actors

The adoption of AI systems for content generation and distribution is occurring alongside the spread of mis- and disinformation. In their response to the draft amendments to IT rules in India, the IT for Change organisation emphasised that “approaches to addressing misinformation and fake news need to be reframed with due cognisance of the information economy and its technological mechanics” (Rajkumar & Ashraf, 2023: 3). The accessibility and sophistication of content produced by GenAI are increasing as these tools provide creative possibilities for producing or altering textual, visual, auditory, and audiovisual data, and are used by both private and state actors (Bontridder & Poulet, 2021).

A survey conducted by Freedom House found that a minimum of 47 countries employed commentators to manipulate online discussions in their favour – double the number of countries involved a decade earlier (Funk et al., 2023). Evidence on how systematic these efforts are, and which specific actors are involved, is missing or weak, even though these mis- and disinformation tactics are growing in sophistication as GenAI tools become more powerful, readily accessible, and user-friendly. It is clear, however, that these tools are being used to foment uncertainty, defame adversaries, and sway public discourse.

False information, propaganda, and hoaxes proliferated with the spread of the Internet and social media. This has increased further with the availability of user-friendly GenAI tools, enabling deepfake creators to build realistic synthetic videos, speech, or images of real individuals without extensive technical expertise or substantial financial resources. For example, CounterCloud – an AI model said to produce automated disinformation that is convincing 90 per cent of the time – was reported to be available at a cost of less than 400 US dollars per month (Funk et al., 2023). This illustrates the cost-effectiveness and simplicity with which significant mis- and disinformation operations can be generated (see Figure 3.1).

Figure 3.1 Example of realistic AI-generated face



SOURCE: Generated using 2020 algorithm StyleGAN2

In the United States, AI-generated information has been used to tarnish the reputations of political rivals. In Venezuela, state-controlled media used AI-generated videos featuring fabricated news anchors from a fictitious international English-language network to disseminate pro-government propaganda. Produced by Synthesia, a company specialising in the creation of personalised deepfakes, this content was widely shared on social media platforms. In 2023, during the Nigerian elections, a modified audio recording created using GenAI was shared on social media. The recording falsely claimed to provide evidence of an opposition presidential candidate's involvement in attempts to manipulate the ballots (Repucci & Slipowitz, 2022; Ryan-Mosley, 2023). A study in sub-Saharan Africa found that people displayed a greater willingness to share mis- and disinformation compared with those in the United States (Wasserman & Madrid-Morales, 2019).

It is sometimes argued that concern about the risks of AI-enabled mis- and disinformation are exaggerated and that it distracts attention from other issues. For example, there is evidence that heavy misinformation consumption is limited to people who are already more likely to seek it out, suggesting that increased information quality is unlikely to have a significant effect (Broniatowski et al., 2023; Motta et al., 2024). It is important, however, to not assume that the results of research on these risks apply globally (Madrid-Morales & Wasserman, 2022).

The impact of GenAI on mis- and disinformation, including deepfake images, can be examined in four categories: increased quantity; increased perceived quality; increased personalisation; and accidental generation of

plausible but false information (Simon et al., 2023) (see Figure 3.2). Measuring the scale of AI generation and the distribution of mis- and disinformation and the impact of mis- and disinformation campaigns is challenging, as discussed in Chapter 1, but evidence does suggest that AI generation and distribution of mis- and disinformation grew significantly in the five years leading up to 2023 (Funk et al., 2023). Empirical studies of bots have concluded that they are omnipresent on social media platforms such as X/Twitter (Keller & Klinger, 2019), although many are used for relatively benign purposes (Makhortykh et al., 2022). And cyber troop (government or political party actors tasked with manipulating public opinion online) activity has been identified in 81 countries (Bradshaw & Howard, 2019).

Figure 3.2 Deepfake image of Donald Trump



SOURCE: Generated using Stable Diffusions

There is also relatively limited evidence on the impact of AI systems on trust in information and news media (see Chapter 2). Sources include incidence reports, corporate case studies, surveys of worldwide campaigns (e.g., the Global Inventory of Organized Social Media Manipulation), and reports on content takedowns by platforms such as Facebook (Windwehr & York, 2020). Meta publishes annual transparency reports documenting its content moderation actions, but these are criticised for not disaggregating the types of content or including the precise quantity of content removed (Bradshaw et al., 2020, 2021). A lack of standards for, and transparency in, data collection makes it difficult to verify and replicate findings (Kostygina et al., 2023).

The United Nations' work on information integrity on digital platforms has documented cases of harm linked to mis- and disinformation, including

violence against individuals and groups ensuing from the posting of hate speech, reduced take-up of Covid-19 vaccination programmes, and risks to the achievement of the Sustainable Development Goals (UN, 2023a). A survey found that 70 per cent of United Nations peacekeepers felt that mis- and disinformation was having a “severe, critical or moderate impact on their work” (UN, 2023a: 13). A synthesis of evidence from 1,300 sources (news articles, academic papers, white papers, and a range of other grey literature) has provided case studies of impact in over 70 countries (Bradshaw et al., 2020, 2021), and studies of the proliferation of mis- and disinformation during the Covid-19 pandemic found a significant impact on vaccine uptake (Fertmann & Kettemann, 2021; Naeem et al., 2021; Posetti & Bontcheva, 2020). Attempts to develop models to simulate the potential impact of mis- and disinformation face challenges even when data access becomes easier, in part, because there are substantial issues to overcome in modelling real data, and many events can affect the impact of exchanges on platforms (Iamnitchi et al., 2023).

AI systems and democracy

Reciprocal relationships between the development and deployment of AI systems and the mediated public sphere, including how these relationships affect news media diversity and media freedom – and more generally, the interaction between these systems and societal resilience and cohesion and social and environmental sustainability – are all important considerations in understanding the impact of AI systems within information ecosystems and on democracy.

Crucially, the use of AI systems for content governance shapes the public sphere in which communication flows occur. While private communication platforms that use these technologies do not themselves set out to directly censor, the design and use of content governance algorithms influences democratic discourses (Elkin-Koren, 2020). Just as AI systems can contribute to more diverse information ecosystems, they can reinforce the monitoring capabilities of authoritarian states and enhance inequalities and unfair power structures through labour extractivism (Adams, 2022; Boix, 2022). Without negating the role of automated tools, it is important to realise that non-technology-related phenomena, such as the quality of a social security system or the degree of support for gender equality, are bigger factors when it comes to furthering societal cohesion and resilience. This means that, in assessing the impact of AI systems, the socioeconomic and political contexts in which information ecosystems operate must be considered together with the policy and regulatory situation of a country or region (Birwe, 2024; Breuer, 2024).

It is important to account for the specific influences that AI systems can have over the composition and functioning of the public sphere. AI tools used by platforms to curate content tend to favour emotionalising content to

increase engagement. Personalisation systems are used to amplify content that emotionalises and divides because platform business models demand engaging content (Bail, 2021; Settle, 2018). This can reward social and political groups that communicate substantially through this content or in a given style (Noble, 2018). Geographically dispersed and fringe groups can profit from easier connections through social media (Kreiss & McGregor, 2023). Social media use is positively correlated with more diverse information consumption in some studies (Gil de Zúñiga et al., 2021; Möller et al., 2018), for example, and the use of interest histories (personalisation based on previous behaviour) to shape information consumption has been found to increase content diversity (Möller et al., 2018). Platforms often receive “outsized attention and criticisms” (Kreiss & McGregor, 2023: 561) as the main drivers of societal polarisation, but broader societal conditions are important since underlying societal inequality can be a bigger threat to societal polarisation. Thus, the relationship between AI systems adoption and societal conditions is reciprocal and characterised by power asymmetries (Kreiss & McGregor, 2023).

The protection of democratic values and supporting the public sphere are key goals of platform regulation (Mökander et al., 2023), and studies of responses to AI systems have an increasing emphasis on user rights-related obligations in court rulings and laws, especially in the European Union (Katzenbach, 2021). Any discussion on the democratic implications of AI systems needs to include Global Majority World voices and develop alternatives to current practices of exercising socioeconomic and geopolitical power through algorithmic tools and datafication (Roche et al., 2023; Ricaurte, 2022). Calling only for transparency of content governance systems that influence global information distribution processes is not enough. Novel approaches to making algorithms used by communication actors more accountable are essential (Ananny & Crawford, 2018), notwithstanding challenges to transparency including information asymmetry, uncertainty, and resourcing. Interdisciplinary engagement and decisions about legal rights to access information, shared decision-making about AI transparency choices, efforts to understand social and societal impacts, and adequate resourcing of transparency teams and audits are all needed (Bates et al., 2023; Ruffo et al., 2023). The extent to which the voices of those in Global Majority World regions will play a role in shaping the rules for AI systems and content moderation is uncertain as countries in the Global North push back on the United Nations’ efforts to give countries in these regions – including China – a strong voice in AI systems governance (Alexander, 2024). However, diverse voices are essential in the light of the United Nations General Assembly conclusion that AI-generated mis- and disinformation could “undermine information integrity and access to information [and] undercut the protection, promotion and enjoyment of human rights and fundamental freedoms” (UN, 2024c: 2–3; see also UN, 2024b).

The existence, and plurality, of independent news media of sufficient quality is impacted by increased use of AI tools for content production which, in turn, is influenced by trends in market concentration triggered by AI systems investment and the impact of algorithmic content production and distribution on media plurality and diversity. The Council of Europe has recommended that automated decision-making processes governing the distribution of online content should “improve the effective exposure of users to the broadest possible diversity of media content online” (Council of Europe, 2018: para. 2.5; see also Heitz et al., 2021). However, measuring media diversity is not a simple task (Ranaivoson et al., 2022). Research suggesting that AI system use in social media has negative effects on content diversity in terms of its distribution may neglect the multidimensionality of diversity (e.g., topic, genre, tone, and plurality). Studies using the concept of exposure diversity – the diversity of information that users see – find that algorithmic personalisation systems have strong positive effects on diversity, partly as a result of the element of surprise. However, highly personalised systems that increase the perceived relevance of specific content for users can reduce the range of information to which users are exposed, although this effect may be reduced by increases in media and information (and AI) literacy (Helberger et al., 2018; Kreps et al., 2022; Möller et al., 2018).

Information ecosystems are connected to other societal systems, and although AI systems are only one factor in societal transformation processes, they can both challenge and enhance societal resilience and cohesion. Societal resilience refers to the ability of a society to react to, and recover from, challenges and disruptions, including short-term disruptions (e.g., armed attacks), medium-term crises (e.g., the Covid-19 pandemic), and long-term challenges (e.g., climate change) (Berger et al., 2023b; Haas & Kettemann, 2024; Kettemann & Lachmayer, 2021; Veale et al., 2023). Societal cohesion is a key contributing factor to, and predictor of, societal resilience. It refers to the capacity and extent to which a society’s members cooperate and work together towards collective well-being based on shared values. These values are shared, questioned, and developed through communication processes.

When automated content moderation tools play an important role in information ecosystems, they can have an impact on societal cohesion, and thus also resilience. Being aware of the rules and practices governing mediated discourse is important for meaningful democratic participation, and increasing algorithmic awareness is an important aspect of AI literacy (de Vivo, 2023; see also Chapter 5). By increasing sensitivity to the impact of AI systems on content production and distribution, societal cohesion and resilience can be better supported. Conversely, research suggests that greater societal resilience is positively correlated with resistance to mis- and disinformation (Kertysova, 2018). Evidence also suggests that use of AI tools by political actors can increase the quality and speed of responses to political queries by citizens,

thus leading to a higher level of engagement, with the important caveat that citizens must be helped to understand and trust how these systems are used (Muñoz, 2023).

Evidence is contradictory regarding whether automated content governance is a main driver of societal polarisation and, hence, of a decline in social cohesion, and polarisation dynamics is a key field of research, as discussed in Chapter 2 (Ruffo et al., 2023). A lack of digital literacy, societal vulnerability towards information “pollution”, and a preexisting fragmentation of society are cited as playing more substantial roles in societal polarisation. The prevalence of mis- and disinformation such as hate speech can lead vulnerable groups to withdraw from online discourses, thus decreasing social cohesion (Breuer, 2024; Nordic Council of Ministers Secretariat, 2023; Sîrbu et al., 2019; Washington, 2023). Although automated content governance systems are not the only contributor to polarisation, news diversity and media consumption practices can clearly be affected by machine gatekeeping (Evans et al., 2023; Ross Arguedas et al., 2022a). Exposure to misinformation and partisan information also can elicit strong emotions, which, in some studies, is shown to lead to some attitude polarisation (Weismueller et al., 2024).

The integration of AI systems into the workplace is also profoundly transforming labour conditions across industries, which impacts social sustainability. Content moderation is essential for maintaining the quality and safety of online platforms but, despite its importance, the job is often outsourced to contract workers who face unstable employment conditions (Ahmad & Greb, 2022). Processes underpinning data collection, content labelling, and training can contribute to harm, including traumatisation associated with working with problematic content or training data, and this affects underpaid workers in countries that lack stringent labour protection laws (Veale et al., 2023).

For example, the Amazon Just Walk Out or Amazon Go stores, where people could “enter at gate, shop and walk out”, employ hundreds of workers in India, and labour issues prompted the company to roll back the use of automated technology in its stores (AWS, 2025). Uber, Lyft, and DoorDash use AI systems and data analytics extensively to manage their operations (Bitter, 2024; Burrell, 2016). Drivers and delivery personnel are typically classified as independent contractors rather than as employees, with many workers not receiving benefits or protections associated with employment, such as health insurance, paid leave, or job security. The intersection of surveillance capabilities, worker monitoring, and labour conditions means that these companies’ uses of AI systems and their approaches to collecting and processing data are attracting attention due to the impacts on worker privacy, autonomy, and rights, with data privacy concerns being raised in regions including Africa (Abdulrauf & Dube, 2024).



In India, the door-step food delivery platform Swiggy gamified insurance for its rider partners. Swiggy's weekly ranking system allows workers to access health insurance depending on the number of "perfect" deliveries they make. Amazon designed a 30-day Delivery Premier League (DPL) for its part-time workers, under the Amazon Flex programme. Modelled after the flagship Indian Premier League cricket event, DPL gamified delivering packages. Each hour spent on the platform collecting packages from mini-warehouses and delivering them to customers' homes constitutes a run – a unit of scoring in cricket. The more hours spent delivering, the more deliveries riders accumulate, ultimately resulting in rewards such as smartphones, motorbikes, televisions, and Amazon gift cards in addition to the flat 125 rupees (about 1.50 US dollars) paid per hour. In many Indian cities, local governments have implemented GPS-based systems to monitor sanitation workers in cities including Patna and Pune, to track the movements of workers to boost productivity and manage schedules, raising concerns about privacy and the dignity of labour. Amazon uses sophisticated systems to track the movements and productivity of warehouse workers who are often required to pack hundreds of boxes per hour, with time spent off-task leading to warnings or job termination (Bansal, 2024; Bitter, 2024; Christopher, 2021; Nagaraj, 2020).



Finally, training state-of-the-art AI models is an energy-intensive process and developments present substantial challenges to environmental sustainability. LLMs demand vast amounts of data, and power-intensive training processes, involving complex calculations, run on thousands of high-powered graphics processing units (GPUs) over several weeks. This can lead to a sizable environmental footprint, because data centres are one of the major drivers of increases in energy demand and greenhouse gas emissions (iea50, 2024). The rapid development and adoption of AI systems is also leading to escalating demands on the digital infrastructure – data centres – that are essential to progress. Goldman Sachs (2024) has predicted that by 2030, data centre power energy demand will grow by 160 per cent, and research suggests that the real emissions from data centres can be more than six times the officially reported values (O'Brien, 2024).

Advances in chip technology can mitigate environmental impacts, offering greater computational outputs per watt of power consumed and thus a potential offset – although relatively minor – to the energy-intensive nature of extensive data operations (Berthelot et al., 2024; Cowls et al., 2023). Google is reported to have used AI systems to enhance the energy efficiency of its

data centres, reducing its cooling energy requirements by up to 40 per cent (Burgess, 2016; Google, 2022). In some countries, resistance to the energy consumption of large data centres and computing resources is emerging and strengthening (Velkova, 2024; WEF, 2024). This ties in with a general growing demand for public participation in decisions impacting sustainability agendas with calls for holistic approaches to these issues, encompassing the whole lifecycle of AI system development, including environmentally responsible innovation (Brevini, 2021; Makan, 2023; Wu et al., 2022).

Conclusion

This chapter opened with a discussion about the misleading use of the term “AI” and the importance of naming specific technologies, such as algorithms or machine learning and LLMs, in research intended to assess the risks for human rights and societal processes of self-determination. We then turned to research on how human rights law applies internationally and at regional and national levels as AI systems, including GenAI, lead to potential and actual rights violations. The analysis focused on AI biases and fairness, freedom of expression, privacy protection, and participatory rights and the issues arising from the uneven spread of AI systems throughout societies, creating a heightened need for states to devise ways of respecting, protecting, and implementing human rights protections.

This chapter also focused on the role of AI systems in managing the visibility and spread of information, including mis- and disinformation – that is, content governance and its impact on information integrity. We emphasised the role of these systems in both content generation and in content moderation and curation, and, especially, the role of algorithmic personalisation systems in contributing to non-transparent decision-making. We examined major changes in how news organisations produce and distribute their content and the capacities of a range of actors to manipulate online information. In all these areas, the importance of transparency and accountability best practices for automatic content governance systems has been emphasised with a view to achieving greater public awareness of how AI systems are becoming embedded within information ecosystems and impacting the integrity of information. We also emphasised that the reciprocal relationships between the development and deployment of AI systems and the broader evolution of information ecosystems is crucially important for the future of the public sphere, societal resilience and cohesion, the social sustainability of labour markets, and environmental sustainability.

Big Tech power and data governance

Introduction

The data that constitutes today's information ecosystems is produced and controlled primarily by a small number of companies in accordance with business models designed to prioritise profit over corporate responsibility for human rights, privacy, and safety (de-Lima Santos, 2023). Decisions by one or more human actors to create a record of something – such as an action performed by a human, usually for a particular purpose – result in the production of this data (Rosenberg, 2013). Although the ecosystem metaphor is helpful imagery for thinking about the dynamic, complex, and interconnected systems that determine who can receive or share what kinds of information and in what contexts, when the metaphor is used to imply that data is a naturally occurring resource, this can distort our understanding of how data aggregation and algorithmic technologies alter the systems through which people build, share, and receive information. Information ecosystems are fundamentally shaped by (and shape) the environments in which they arise – in today's context, a data-rich environment.

In this chapter, we bring a political economy perspective to bear on the way power relations establish the conditions for how data economies operate, and how they are inequitably experienced by people around the world (Aaronson, 2021; Aguerre et al., 2024; Padovani et al., 2024; Taylor et al., 2022; Verhulst & Schüür, 2023). The analysis draws upon research that critiques data governance measures that leave the business models and data practices of Big Tech companies largely in place; models that incentivise these companies to turn a “blind eye” to mis- and disinformation and which lead to injustice.

Information ecosystem contestations

Big Tech control strategies disadvantage populations in both the Global Majority World and in the Global North by favouring approaches to governing data that produce or entrench social injustices, including wealth disparities and racial oppression. These strategies support data aggregation and AI systems that are changing the way people build, share, and receive knowledge.

This is achieved by exerting monopolistic control through data extractivism (the large-scale harvesting of data by private companies), which leads to dependencies of individuals, communities, and industry sectors on the Big Tech companies. Lurking behind the proliferation of mis- and disinformation (and other kinds of low-quality information), and the reasons that useful and high-quality information is often inaccessible, is the problem of data governance that is not designed to serve a vision – arrived at democratically – of how data should, or should not be, collected and used. These strategies work against information ecosystems that are consistent with democracy; they interfere with political deliberation that is essential for the survival of participatory democracy. Data governance is needed to encourage data and digital infrastructures that enable political communities and allow them to shape the affordances of information ecosystems.

The OECD (2022a: 13) defines data governance as follows:

The diverse arrangements, including technical, policy, regulatory and institutional provisions, that affect data and their creation, collection, storage, use, protection, access, sharing and deletion, including across policy domains and organisational and national borders. Efforts to govern data take many forms. They often seek to maximise the benefits from data, while addressing related risks and challenges, including to rights and interests.

This definition signals the importance of ensuring that datafication processes embedded in information ecosystems protect human rights, but it is agnostic about the monopolistic activities of Big Tech companies, such as practices for data capture and control. Much academic research and civil society advocacy examining contestations around the dysfunction of today's information ecosystems focuses on AI systems and algorithms. However, to think creatively about what it will take to build just and equitable information ecosystems, it is essential to examine the fundamental problem of how to govern the creation of, access to, and use of massive amounts of data. Insight into contestations around data governance, including how data is used in today's information ecosystems and the prospects for supporting – or even enhancing – democratic governance, comes from sources in the critical literature.

In the 1950s, Hannah Arendt worried about a near future in which human technology would replace human thinking. In *The Human Condition*, she worried not because machines would become intelligent, but because of the many things machines would make it possible for human beings to do *without* intelligence:

It could be that we, who are earth-bound creatures and have begun to act as though we are dwellers of the universe, will forever be unable to understand, that is, to think and speak about the things which nevertheless we are able to do. In this case, it would be as though our brain, which

constitutes the physical, material condition of our thoughts, were unable to follow what we do, so that from now on we would indeed need *artificial machines* to do our thinking and speaking. If it should turn out to be true that knowledge (in the sense of know-how) and thought have parted company for good, then we would indeed become the helpless slaves, not so much of our machines as of our know-how, thoughtless creatures at the mercy of every gadget which is technically possible, no matter how murderous it is [emphasis added]. (Arendt, 1958/1998: 3)

One way Arendt's fear has been realised is by creating infrastructures for our social, political, cultural, and economic systems that are pervasive, determinative, and invisible (Cohen, 2023; Star & Ruhleder, 1996). Digital infrastructure that enables information ecosystems and which shapes the integrity of information is not disembodied; it involves familiar material forms of industrial infrastructure including "data centres distributed throughout the world and made up of servers, routers, switches, and miles of cables, as well as redundant power sources, cooling and ventilation systems, and security apparatus" (Sacacas, 2021: para. 7). This infrastructure is just as transfiguring of physical landscapes as railroads, highways, and power grids are. The difference is that the digital infrastructures that "shape, enable and sometimes deliberately constrain life in common" (Stiefel et al., 2024: para. 3) are largely hidden from the conscious experience of the people who depend upon them for many aspects of their lives. The more that we depend on this infrastructure – to get from place to place, to shop, to access government services, to work or go to school, to get medical care, and to have private conversations with friends – the more we become part of it. The more we become part of this infrastructure, the less we are aware of it, and the more it then shapes our perceptions of everything (Barba-Kay, 2023; McLuhan, 1964; Morozov, 2013; Zuboff, 2019).

The risks of treating the digital environment as if it is natural vary depending on the context. There may be few downsides to obeying traffic lights within a well-designed algorithmic road safety system; however, a variety of risks is associated with algorithmic traffic management systems, such as privacy infringement, especially when system maintenance depends on multilayer real-time surveillance (Local Progress, 2024). The consequences of the use of digital infrastructure structures and our reliance on data are also politically significant (Herman & Chomsky, 1989), especially when these systems are involved in creating, sharing, and disseminating information.

Truly democratic data governance requires that it be feasible for polities to contest the design of these infrastructural systems and the means for controlling the uses of the data that makes information ecosystems possible. Digital data-dependent systems increasingly define and constrain political discourse and activity in many contexts in the Global North. For those regions of the world that have not had the opportunity for substantial input into

shaping the data economy, it is crucial to acknowledge parallel, and often conflicting, visions for governing the generation and uses of data (Abdulrauf & Dube, 2024). It is also essential to recognise intersectional perspectives, including gender-sensitive approaches to data governance, and how they reveal various dimensions of bias, discrimination, or disenfranchisement (Chair, 2024), and help to resist the common assumption that systems for creating and sharing information must rely on digital data.

The ubiquity of this assumption is unsurprising, since digital infrastructures are established globally, and this has occurred largely outside deliberative democratic processes. For example, no local, national, or international legislative body considered whether it would be a good idea to set up a system whereby people access news through personalised digital filters designed to maximise the likelihood that they will eventually buy something. Digital infrastructures have been imposed primarily because of under- or unregulated corporate activity operating alongside opaque government procurement processes in the case of provision of services (Calo & Citron, 2021; Colclough, 2022; Crump, 2016; Hardy & Williams, 2008; Zuboff, 2019). Many factors have helped to create the conditions in which the data-related features of digital infrastructures are proliferating, typically with little political friction. None is more crucial than the lack of robust, and robustly enforced, rules about which public and private actors can do what with respect to data (Cohen, 2019).

Political friction and contestation can, however, generate political participation and change (Gordon-Tapiero et al., 2023; Salehi et al., 2015). Thanks to public concern about corporate data practices that followed OpenAI's release of ChatGPT and the ongoing generative AI (GenAI) "arms race", discussion about data governance is at an all-time high. Data governance – in relation to the privacy, security, and the integrity of data – has been addressed historically by policymakers, but political pressure is now strong from within civil society to think about data governance as a lever for restructuring the markets in which Big Tech companies operate. This pressure is galvanising efforts to protect people against human rights infringements and against concentrations of power and wealth that result in datafication practices that are inconsistent with democracy (Giblin & Doctorow, 2022; Mejias & Couldry, 2024).

This gives us an opportunity to question the roles of digital data, data-dependent digital infrastructures, data markets, and companies in the data business in the very formation and function of information ecosystems. Questioning of this kind must be part of any democratic digital-policy project, but any such project must also preserve and promote the capacities of diverse communities to take up such questioning *outside* formal policymaking spaces. Such questioning is necessary not only *for* democracy, but *as* democracy (Benson, 2019; Chambers, 2023).

What is at stake when data and information about the world is structured by technologies that few understand, and even fewer control, is not so much

the ability to resist the manipulations of technologies (as important as that may be); consistent with Arendt's view, it is the ability to think and deliberate with others about the meaning of information and of information ecosystems in relation to the "common good" (Mazzucato, 2023). The issues here extend far beyond protecting and promoting an inclusive public sphere, because the data practices that undergird information ecosystems have profound social, economic, and political implications, including environmental impacts and changes in wealth distribution).

Corporate data monopolisation and digital infrastructures

For people living in places with a highly developed digital infrastructure, it is almost impossible to live without creating a digital record of their lives. This is increasingly so in the Global Majority World as this infrastructure starts to become more accessible in these countries. The publication of *The Age of Surveillance Capitalism* (Zuboff, 2019) coincided with a surge of attention to the ways that corporations track, record, and analyse our online activities to predict and shape consumer behaviour. This is the data that companies such as Google, X, Meta, and Microsoft collect when people use their apps and services to create, find, consume, and share information. This is only a small part of a vast data surveillance landscape that includes most of the government bureaucracies and private entities that manage services and goods industries crucial for the public.

In a data economy, data is produced about us. This happens when we take public transport or drive on public roads, when we work at our jobs, when we open a bank account or apply for a credit card, when we have an interaction with a police officer or seek judicial intervention, when we apply for public benefits, when we rent or buy a home or sign up for electricity, when we go to school, when we go to the doctor, when we interact in online spaces, and so on. As the amount of data and the number of data repositories grow exponentially, so do networks and digital processes for sharing and selling data (see Díaz & Levinson-Waldman, 2020; Ajunwa et al., 2017; Brayne, 2020; Eubanks, 2018; Harwell, 2021; Hooper et al., 2022; Ledford, 2019; Corrales Compagnucci et al., 2022).

Government agencies are often unaware of who has access to the data they produce about their constituents (Harwell, 2019), even if there are rules about what the data collector can do with it, with whom they can share it, and what a third party can then do with it. Most data governance frameworks consist of "a patchwork of national regulatory regimes, multilateral bodies, corporate policies, and multi-stakeholder organizations" (LaForge & Gruver, 2023: 32), which have not proven sufficient to protect most kinds of data. The result is that much data is acquired by companies that use it to generate profits or to

amass power (Lamdan, 2022; Mulligan & Godsiff, 2023; Zuboff, 2019). The political economy of this data infrastructure is one in which almost everyone is comprehensively surveilled. The companies whose products and services enable this are monopolising this data, and the risks and burdens are falling unequally on different groups, weighing most heavily on individuals who are already economically disadvantaged, vulnerable, exploited, marginalised, or already targeted outside the digital context or subject to any form of group oppression. Data extraction without, or with weak, consent processes and biased algorithms, and deployed in ways that benefit corporate economic performance, promote these datafication practices (see Benjamin, 2019; Browne, 2015; Fontes et al., 2022; Graham & Dittus, 2022; Noble, 2018; O’Neil, 2016).

The more companies achieve control of data, the more difficult it is to enact structural and systemic changes to address injustice and inequality. A wide variety of corporate data practices contributes to dysfunction and unfairness. Most of these involve two types of monopolistic activity: monopolisation of user data (i.e., all the data produced about us), which makes money for companies by converting information seekers into information products offered for sale to advertisers; and monopolisation of knowledge (i.e., data organised as usable insight) and information that makes money by converting data resources (including public data resources) into private assets. These pervasive forms of datafication give rise to numerous forms of data dependency.

Data monopolisation and data dependency

In 1998, Larry Page and Sergey Brin were concerned about how advertising revenue might affect the integrity of their newly launched Internet search engine, Google: “We believe the issue of advertising causes enough mixed incentives that it is crucial to have a competitive search engine that is transparent and in the academic realm” (Brin & Page, 1998/2012: 3832). Three years later, after unsuccessful attempts to sell Google to other search companies, they began selling advertising based on user data. Between 2001 and 2003, Google’s revenue increased 3,590 per cent, from 19 million to 3.2 billion US dollars (Veliz, 2021). In the Global North, four companies – Alphabet (Google), Meta (Facebook), Amazon, and Microsoft – now largely control people’s experience of using the Internet to discover and share information. The logic of the advertising attention or “eyeball” economy dictates the kinds of information a person can find or receive, whether that information relates to spring fashion trends, the next election, or the history of capitalism. These Big Tech companies also dominate advertising markets: Facebook and Google together control 70 per cent of the market in the United States and over 65 per cent in the United Kingdom (Doctorow & Giblin, 2023). The troves of data they collect about their users in all the countries in which they operate have enabled them to attain this market dominance,

and other digital platform companies operate in a similar way, even if they command smaller market shares.

The Big Tech companies do not limit their data collection activities to the data that they extract about people using their digital products; they also buy or license data from other companies (and acquire data analytics companies) that gather data from a wide range of public and private sources, and they scrape and aggregate massive amounts of data from every corner of the Internet (Lamdan, 2022; Savitz, 2019). These companies are not transparent about their uses of the data they purchase from third parties or compile from public sources, using it for targeted advertising and product development – including the training of large language models (LLMs) and the algorithms that structure the information flows that drive consumer activity. They are known to sell their own data in ways that are extremely dangerous for democratic societies when it leads to political abuses when governments or their agents, such as police forces, demand access to individuals' non-public data (Biddle, 2024). If, for example, hospital systems become dependent on a managed care algorithm owned by Microsoft, the company would have significant leverage over hospital decisions about how to deliver medical care, and the company could make it difficult for governments to limit their data practices and data hoarding since the company requires data to train and maintain the technology undergirding the managed care algorithm (Tucker, 2023).

The fact that these companies take data from the Internet without having to justify or compensate the data owners is not the result of any affirmative policy to permit such activities. In many cases, despite data-protection legislation, it is the companies that determine what, if any, limits to data extraction they will abide by, and they can change these at any time. An example is Google's 2007 terms of service:

You give Google a perpetual, irrevocable, worldwide, royalty-free, and non-exclusive license to reproduce, adapt, modify, translate, publish, publicly perform, publicly display and distribute any Content which you submit, post or display on or through, the Services. (cited in Mejias & Couldry, 2024: 12)

In the absence of robust data governance, tech companies treat data as an exploitable resource and, following a playbook similar to the history of European colonialism, use that data to create conditions in which resisting their continued use of that data becomes both difficult and costly (Mejias & Couldry, 2024). Big Tech companies use their power to amass data to reinforce their advertising dominance, squeezing out competitors and making it difficult to develop a product or service around an alternative set of data practices, or to use, test, and scale up a framework for information sharing that does not depend on advertising.

Data monopolisation is becoming more extreme as companies race to acquire the massive amounts of data needed to train machine learning models

to perform sophisticated classification tasks and predictive modelling, as discussed in Chapter 3. These technologies, marketed as “AI”, are extremely expensive to build because of the amount of data and computer processing power required for training (Mulligan & Godsiff, 2023). This means that only a small number of companies have the resources to compete in the development of data-intensive algorithms, and they are betting on an eventual payday that will justify their exorbitant initial investments (Metz, 2023; Novet, 2023). The Big Tech companies are integrating these technologies into their models to generate advertising revenue, for example, using GenAI chatbots to mediate search activity. It is unlikely, however, that advertising revenue alone will yield profits sufficient to justify the size of these companies’ investment in data-intensive technologies.

Every indication from Big Tech marketing and public relations documents is that their plan is to develop sophisticated industry-specific digital products that will claim to improve efficiency and reduce costs, while creating path dependencies that lock in their client companies to their algorithmic products. Bill Gates (2023), whose company has invested billions in OpenAI (ChatGPT), has predicted that AI systems use will redefine whole sectors of the economy. Following the example of enterprise resource planning system providers such as SAP promising efficiency gains, OpenAI highlights “AI” products for sectors including healthcare and legal services (Ven et al., 2008; Ferràs-Hernández et al., 2023; Melih, 2022; van der Vlist et al., 2024).

The government in China is facilitating a Big Tech oligopoly which seeks to rival the economic and political power of the largest United States-based tech companies. Its national technology champions are strategically cultivated – partly by banning foreign competitors and partly by using policy incentives to favour domestic firms (Tusikov, 2021). Large tech companies in China control digital platforms for social media, e-commerce, search, and online payments (Borgogno & Zangrandi, 2024; Tusikov, 2021; Wang, 2023), and the country is expanding its geopolitical influence through the Digital Silk Road initiative (Erie & Streinz, 2021). This supports the export of Chinese information and communication goods and services as well as technical standards to countries, especially in Africa and the Indo-Pacific, where investment in digital infrastructure is at a stage where dependencies can be created:

While it is true that China has consistently advocated for national autonomy over data governance issues, it is also trying to establish technological and infrastructural dependencies within the digital space of several countries. Technological dependences are established whenever the digital infrastructure relies on standards, software and hardware that cannot be maintained without active support from Chinese players. (Borgogno & Zangrandi, 2024: 19)

The problem with these Big Tech strategies is not just that a small number of companies control most of the world's digital data or that they are using it primarily for self-enrichment. It is also that weakly governed corporate practices are what determined that digital data is being produced on such a scale and commercialised by default (see Cohen, 2019; Melih, 2022; Mulligan & Godsiff, 2023; Sadowski, 2019). It is this profoundly undemocratic, economic fiat that has successfully preempted meaningful political deliberation about rights to data ownership, what role data should have in the private and public sectors, how it should inform bureaucratic decision-making, and whether and in what contexts data extraction should be minimised or prohibited. While there are conversations about this in some political communities, the terms of debate around government initiatives to legislate the curtailment of Big Tech practice are limited by the lobbying of companies and business associations across industries that have already become entrenched in data dependency relationships (Rankin, 2023).

The burden of having ceded so much ground to tech companies is high. When it is assumed that systems for extracting and disseminating information must operate by reducing information to codable data and use AI systems to determine how data should flow, it becomes extremely difficult to ask questions about what a *good* system should look like. It may be possible for governments to compel tech companies to deliver technologies that deploy algorithms that are more accurate, useful, inclusive, or accessible; however, fixing the operational flaws of algorithmic systems is ultimately a trivial problem when compared with the problem for democratic society when it is not open to citizens, their representatives, or the wider political community, including migrants and refugees, to contest the design and function, and even the existence, of these systems. This kind of thinking involves questioning the premise – and it is fundamental to political deliberation in democratic societies. Ironically, even the most sophisticated AI systems question their own logics, that is, reflect on whether the learning being done is the *right* learning for the problem at hand (Green, 2022; Green & Kak, 2021; Selbst et al., 2019).

Business models, knowledge, and mis- and disinformation

Big Tech platform business models make mis- and disinformation-based campaigns an attractive option (Bontcheva et al., 2020), as discussed in Chapter 2. A core driver behind these models is an economics of emotion that depends on attention and incentivises the creation of mis- and disinformation: “As the design of the algorithms and interfaces of globally dominant social media platforms maximize emotional engagement, we regard social media as a primary site of datafied emotion worldwide” (Bakir & McStay, 2022: 32). Digital influence mercenaries exploit platform affordances on behalf of their clients, and mis- and disinformation are used to gain platform users' attention,

transforming this into a commodity for sale to advertisers. The introduction of the “like” button was a key step in the evolution of platform affordances that served the interests of influence mercenaries and the Big Tech companies (Bakir & McStay, 2022). This button gave Facebook (now Meta) a huge new source of valuable data about its users by tapping into their feelings, and this enhanced its personalisation offering to advertisers. Meta and other platform companies continually test algorithms on their users, for example, using “A/B” testing, where the reactions of users to different versions of a website are tested, often without the users being aware.

Almost all platforms have adopted variations of vanity metrics, using them to algorithmically curate content posted by other users to identify specific users as recipients and as targets for advertisers (Rogers, 2018). How this content curation is performed varies among platforms and is not to be confused with content moderation, which assesses whether a post is acceptable under a particular platform’s rules (see Chapters 2 and 3). On TikTok, user “likes” are combined with their content views (Benton, 2022), and 4chan curates user posts so that only the most liked ones survive, which leads to the promotion of more extreme material (Tuters & Hagen, 2020). Platforms such as YouTube reward content producers, and this incentivises the creation of more extreme content, and X has adopted a variation of this policy (Pequeño, 2023).

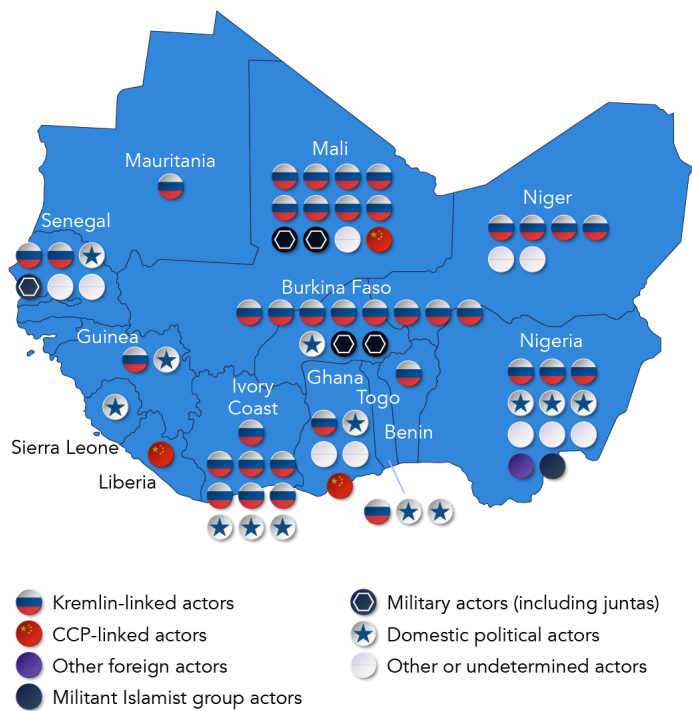
The question is whether these business models *inevitably* lead Big Tech companies to turn a blind eye to mis- and disinformation. The reason the attention economy is key to platforms’ financial viability is due to the advertising revenues that user attention helps generate. Extreme content increases user engagement: the economics of emotion monetises deception online, first, via a service contract with digital influence mercenaries to exploit platform affordances to achieve a client’s strategic objectives, and second, by attracting user attention through deceptive content and selling this attention to advertisers (Bakir & McStay, 2022). Having no content moderation policies (or policies that are not implemented) can be bad for business, as illustrated when advertisers terminate their business with platforms when their ads have appeared alongside offensive posts, such as hate speech, triggering brand boycotts (Zhu, 2022). On balance, however, platform business models incentivise a very light-touch approach to content moderation. Access to facilities such as platform application programming interfaces (APIs) provide a means to programmatically interface with platform data to harvest millions of user profiles that are then used to target political advertising (Hinds et al., 2020; Gorwa & Guilbeault, 2020; Zammit et al., 2021). Some platforms such as X have had very permissive policies, allowing developers to use APIs to create bots that automatically tweet. Others such as Meta impose tighter controls, but the management of the use of “sock puppet” accounts (i.e., false identities used for purposes of deception) is a challenge.

Sophisticated ways of using AI systems to create platform affordances enable exploitation of the capacity of controversial content to capture user

engagement. These affordances are being exploited by influencers who build a following and channel content that is designed to sustain their followers’ engagement, which they then monetise (Diaz Ruiz, 2023). A market-shaping approach helps to reveal how market-makers bring markets into existence by incentivising the generation of viral content – and “clickbait” – which boosts the circulation of “controversial claims, adversarial narratives, and deceptive content”; in this context, mis- and disinformation are an “expected outcome, not breakage” of a platformised media market (Diaz Ruiz, 2023: 2188).

Political campaigns world-wide are increasingly data-driven as the platforms’ capacities to deliver targeted advertising become more sophisticated and policies regarding transparency are absent or unenforced (Mehta & Erickson, 2022). The Africa Center for Strategic Studies reported that mis- and disinformation campaigns increased nearly fourfold from 2022 to 2024, with a total of 189, nearly 60 per cent, sponsored by foreign states (Russia, China, and the United Arab Emirates were prominent) (Africa Center for Strategic Studies, 2024) (see Figure 4.1).

Figure 4.1 Map of mis- and disinformation campaigns in West Africa, and state sponsorship



SOURCE: Africa Center for Strategic Studies, 2024; redesigned by Matthieu Mercier and adapted by Nordicom.

Responding to concerns about political advertising's lack of transparency, in 2018, Google, Facebook, and X/Twitter established political ad archives, including information about advertisers. However, various factors, including doubts about the capacity of citizens to find and understand data and financial incentives, work against the delivery of transparency in an effective and meaningful way (Mehta & Erickson, 2022). The United Nations has argued that "digital platforms should move away from business models that prioritise engagement above human rights, privacy and safety" (UN, 2023a: 23). A report commissioned by NATO's Strategic Communications Centre of Excellence found that "buying manipulation remains cheap" and "the gap between countering inauthentic engagement and platform reporting is widening" as platforms focus more on limiting the reach and impact of messaging and less on denying access to commercial information manipulators (Fredheim et al., 2023: 3). However, the cost to influencers of buying manipulation has been stable over time (Bradshaw et al., 2021).

Those intent on spreading mis- and disinformation exploit the platforms' business models, and the platforms are complicit in campaign propagation (Posetti & Bontcheva, 2020), providing a target for cyber troop activity (i.e., government or political party actors tasked with manipulating public opinion), with activity identified in at least 81 countries (Bradshaw & Howard, 2019). With anonymity accorded to online labourers who participate in these activities (Shleifer, 2004), it is unsurprising that 87 per cent of participants in an experiment were found to be willing to accept jobs involving the creation of mis- and disinformation (Cohn et al., 2022). In a highly competitive labour market, people persuade themselves that if they do not take on work, others will.

Data governance rules permit and even foster the amplification of mis- and disinformation through coordinated influence operations. In Venezuela, an interview-based study revealed a range of influencer motivations, organisations, technical systems, adversaries, and strategies, including recruiting and paying influencers with campaigns organised through both hierarchies and decentralised operations. Examining influencers' incentives to contribute, their organisational structures, resources, capabilities, and limitations, the strategies and operations they employed, their views of disinformation, platform defences, and their efforts to evade and recover from detection, those influencers who were propagating mis- and disinformation were found to learn how to evade platform defences (in this case, X's) (Recabbaren et al., 2023).

Other studies emphasise the political economy of this activity – and not just the roles of individual actors. A study of mis- and disinformation campaigns in the Philippines and Indonesia concluded that legacy media's history of ownership and political collusion in postcolonial societies make them vulnerable to narratives about bias and bigotry. This research highlighted "the broader (Western) discourse that has positioned these two countries as examples of global South contexts where social media have 'ruined democracy', insofar

as masses of voters are assumed to have been duped by digital disinformation campaigns” (Ong & Tapsell, 2022: 252). These features of the data economy are relatively invisible unless the political economy of infrastructures and data monetisation is investigated to show how power relations and monopolisation strategies enable platform complicity in the villainy of mis- and disinformation actors (Ong & Tapsell, 2022).

When the business models of pre-digital broadcasting media, partisan media, and digital media platforms are compared, they are associated with different forms of mis- and disinformation. In the pre-digital media era, offending the audience was often considered bad for business and, in some countries, this encouraged the presentation of information in ways that aligned with the beliefs and values of the majority, with fewer tendencies towards the proliferation of mis- and disinformation. With the growth in opportunities for the emergence of online partisan media, confrontational narratives and the promotion of minority viewpoints have become common, while partisan or state-controlled media are present in some countries. The question is how it might be made more costly for the major generators of mis- and disinformation to engage in this activity. At present, the powerful Big Tech companies’ business models are counterproductive to efforts to tackle mis- and disinformation. This means that greater attention needs to be given to creative approaches to strengthening data governance and democratic decision-making about who should be able to make use of data (see Chapter 6).

Big Tech monopolisation of data and AI system biases

Advances in AI systems and their integration into search engines and conversational applications such as ChatGPT yield both benefits and harms related to algorithmic bias and information manipulation that impact people’s decisions and societal norms (see also Chapter 3). The AI industry depends on data extractivism to build its systems and is dominated by powerful companies (mainly in China and the United States). This concentration of power is supported by disproportionate access to resources, including computing power, high-quality datasets, and expert talent, which disadvantages smaller entities and independent innovators since these resources are often beyond their reach (Luchs, 2023). Talented AI professionals are drawn to large firms offering better financial and career incentives, resulting in talent consolidation and excluding smaller players and many academic institutions (MIT Technology Review Insights, 2023). The high cost of entry and not having access to proprietary datasets disadvantages those seeking to develop supportive community data governance frameworks and practices.

AI models, including those powering GenAI, incorporate biases from their training data and can develop new biases through interactions with their users (Ferrara, 2024a, b). These biases manifest in several harmful ways, including the reproduction of existing biases and through their impact on user

perceptions. Just as biases are present in human-generated content, LLMs can inadvertently perpetuate these biases, leading to a skewed representation of facts and socially constructed biases in AI-generated content. These biases can then shape how individuals perceive reality, reinforce stereotypes, generate mis- and disinformation, and present a biased view of events and histories; they are an opaque form of epistemic privileging (Machill, 2020). If an AI system is trained on historically biased data, it may generate responses that are subtly prejudiced, reinforcing harmful norms instead of challenging them.

Until recently, Internet users conducting information discovery using search engines such as Google, Yahoo, or Bing have received results as a list of sources ranked according to relevance criteria defined by the search engine, which is influenced by advertising expenditure, not simply by counting the number of in-bound links to a web page. Users can still choose which links to follow, but a search engine manipulation effect is found to influence consumer choices, and even voting behaviour, when search result rankings are manipulated or result visibility is changed (Epstein & Robertson, 2015). GenAI is now being used to create conversational search assistants that summarise search results instead of listing them (e.g., Microsoft's Copilot integrating ChatGPT into its Bing search engine and Google's Gemini search tool). This may reduce a user's ability to discover diverse viewpoints, potentially limiting exposure to multiple viewpoints, diminishing abilities for critical thinking and source evaluation skills, and decreasing agency over the information that is consumed (Hadi Mogavi et al., 2024).

GenAI systems may unintentionally misinform users due to hallucinations (plausible responses that have no basis in reality) that are a consequence of the underlying technology (Bender et al., 2021; Xu et al., 2024). When GenAI systems are used to target certain users with specific content, this has the potential to subtly influence opinions and behaviours, potentially unethically (Motlagh et al., 2023). This can lead to confusion and the spread of mis- or disinformation, and the risk of deliberate or voluntary manipulation is argued to require the introduction of guardrails (Hao et al., 2024; Vasudevan et al., 2024).

Bias is an element of human cognition that can serve as a heuristic for faster decision-making in complex environments (Kahneman et al., 2021). When embedded within AI systems, however, biases can perpetuate harm, especially when they inflict unfavourable outcomes on individuals or groups as a result of inaccuracies and distortions in decision-making. The potentially substantial consequences of decision-making based on flawed or biased data include financial loss, reputational damage, and legal penalties. Bias can manifest as direct discrimination when AI systems provide less favourable results to users from certain demographic groups (Angwin et al., 2016). When these systems are used as the decision-maker for social justice programmes, this can lead to the exclusion of marginalised groups, and they have been shown to worsen existing inequalities and diminish trust in data-driven systems

(Park & Humphry, 2019). Similarly, hiring algorithms can show preference for candidates of a specific gender or racial background irrespective of their qualifications (Raghavan et al., 2020).

All biases cannot be eliminated, and the risk of unfavourable treatment of individuals or groups persists. To mitigate risk, organisations need to implement comprehensive data governance frameworks with clear guidelines for data quality, usage, and security (Marcucci et al., 2023). Examples are frameworks that include standardised procedures for data collection, validation, and storage, and that suggest software tools to manage data quality and correct inaccuracies in real time (Veiga et al., 2017). Regular bias audits and algorithm reviews are crucial to minimise risks. These can be facilitated by third-party auditors, and fairness tools can be used in machine learning to adjust LLMs (Verhulst, 2024). Data governance frameworks and practices can help to improve data integration and address biases when they adopt transparent data architecture principles, and some enterprise data management platforms can make it easier to handle large datasets. By understanding the sources and impacts of data analytics flaws and employing effective mitigation strategies, organisations acquire capacities to improve data integrity and decision-making accuracy. This is essential if AI systems are to be used responsibly and ethically to promote equity in automated environments (Mitchell et al., 2021).

Bureaucratic oversight mechanisms like auditing and litigation, however, have had a limited impact on the power wielded by Big Tech companies. Even when regulation creates new enforcement mechanisms that can be used to compel transparency concerning the corporate management of data and to impose penalties on companies that fail to comply, all actors do not have the resources or political will to use these mechanisms in a way that substantially changes the Big Tech companies' monopolistic data practices. The European Commission's task force with responsibility to "check whether some of the world's richest tech firms are complying with rules designed to make them cede some ground to their smaller competitors" is under-resourced (Hancock, 2024: para. 2). In the United States, there is no federal legal framework regulating corporate data practices, but there are legal frameworks relating to fair labour, competition, intellectual property, and contracts that could be used to challenge the dominance of the Big Tech companies. Until recently, however, there has been little sign of significant enforcement activity against tech companies in the United States, and when cases are brought and even won, they involve lengthy proceedings (Landau, 2020). In summary, AI systems offer unprecedented access to information and the ability to analyse vast datasets, but they require careful management to mitigate risks. Developers and policymakers must collaborate to implement robust ethical guidelines and transparency measures if advances in AI systems are to contribute positively to society and without compromising the integrity of information.

Monopolising knowledge by foreclosing access to information

In parallel with the biggest tech companies in the Global North, a cohort of lesser known, but equally powerful, companies operates to make money by amassing and analysing and then selling analysed datasets to other companies and institutions. Examples of powerful data analytics companies in Western countries are RELX, Thomson Reuters, and Experian, which hoard and sell raw datasets, and – more importantly – information that has already been extracted from them (Lamdan, 2022). RELX and Thomson Reuters operate as a duopoly with respect to legal information in the United States, owning the only two robust databases for conducting legal research, both with high subscription fees. Experian dominates the international data analytics market for financial information, based in 32 countries and collecting and analysing data from thousands of sources, and selling it with tools for integrating digital information to businesses to support clients in more than 100 countries. Smaller specialist data analytics companies serve specific sectors, but the trend is towards the large companies buying others and becoming behemoths across all markets (Gautier & Lamesch, 2021; Lamdan, 2022; Larivière et al., 2015).

Some purchasers of their digital products are other Big Tech companies, such as Google in the search market and the Big Tech companies that incorporate the products into AI systems for search, newsfeeds, or ad placement on their platforms. A large proportion of the customer base for these data analytics companies is government agencies, financial institutions, law firms, universities, healthcare companies, and legacy media and news organisations. These entities can typically pay for the products that data analytics companies sell, which positions them as economic, social, and political gatekeepers within society. The result is that an enormous amount of information with profound public interest value is removed from the public sphere and reconstituted as the intellectual property of companies (Larivière et al., 2015). Although some government agencies provide Open Access to information with the costs of acquiring and curating it borne through taxation, they increasingly outsource to the services offered by companies. A person who wants to access, for example, a newspaper article in the archive owned by RELX containing “5 billion documents and records from over 35,000 sources” of local and international news (LexisNexis, 2024), must either be wealthy enough to afford a personal subscription or be affiliated with an organisation that has one. The consequence of placing so much information behind paywalls is often tragically concrete: “Doctors battling malaria outbreaks in Africa can’t read reports about lifesaving medications and measures. They can’t afford to read past the articles’ abstracts” (Lamdan, 2022: 53).

The situation is especially problematic when it comes to academic research, much of which is publicly funded but very little of which is available to taxpayers who do not belong to an elite institution (Demeter, 2019; Harvie et al., 2013; Nettle, 2023; Puehringer et al., 2021).



Elsevier was founded in 1880 as a publisher of scientific and medical research. Now owned by RELX, it is dominant in a group of five companies that control access to academic information products globally. Elsevier publishes over 500,000 articles annually in 2,500 journals, and its archives contain over 17 million documents. Elsevier charges universities exorbitant fees for access to its journals, packaging them to boost profit margins, rather than to serve the needs of libraries and the people who use them. Elsevier makes thousands of dollars from journal articles costing an average of about 600 US dollars each to produce. Elsevier and other companies in the scientific, technical, and medical unit generated an operating profit of 38 per cent in 2023 (compared with Walmart's 3% and Toyota's 12%). As an information analytics business, Elsevier harvests data from its own content stores to create digital products that score and rank universities, journals, and scholars according to prestige and influence metrics. Its services are used to predict which research projects are likely to be successful, academic funders use them to direct financial support, and these products may be used to make hiring decisions in ways that do not support the work of researchers, students, or librarians, but rather the money-making interests of research institutions (Nicholson, 2024; The Majority Report w/ Sam Seder, 2023).



Some “seventy-five percent of academic research is paywalled, and it usually costs around \$30 to look at a single journal article” (Lamdan, 2022: 53). These figures apply in the United States and charges are even higher in other parts of the world. This is changing, with many academic funders in the Global North mandating that researchers publish papers (and data where available) under Open Access rules (EC, 2016a; NSF, 2023), but this requires that a university or individual pay the Article Processing Charge costs in most cases. In addition to creating access-to-knowledge disparities based on institutional privilege, this creates disparities among institutions. Universities in the Global Majority World are less likely to have the resources to purchase subscription services within which companies like RELX trap the academic articles they own. For instance,

in 2008, Harvard subscribed to 98,900 serials and Yale to 73,900. The best-funded research library in India, at the Indian Institute of Science, subscribed to 10,600. Several sub-Saharan African university libraries subscribed to zero, offering their patrons access to no conventional journals except those donated by publishers. (Suber, 2012: 30; see also Peters, 2016)

Scientists from all over the world recruited 27 established institutions to try to access fulltext paywalled articles in the field of ophthalmology. The results showed that at 15 of those institutions, researchers could access less than half the articles. Those from institutions in wealthier countries (e.g., the United States and the Netherlands) were more likely to access most of the articles. Those at institutions in Pakistan and Ecuador were unable to access any of the articles (Boudry et al., 2019). This kind of control exercised through monopolisation strategies in markets for academic research contributes to huge disparities in the information ecosystem, and inequalities in knowledge production are likely to escalate as publishers such as Elsevier, Wiley, and Taylor & Francis acquire Open Access repositories (Posada & Chen, 2018).

Research independence and verifiable knowledge

Knowledge creation within information ecosystems is also subject to bias when what counts as verifiable knowledge is questioned and scholars are pressured to retract or revise their conclusions by journal editors (Teixeira da Silva, 2021). The independence of research is challenged when results are deemed politically sensitive or they question company or government claims, which leads to overt or subtle silencing of researchers. In some countries,

scholars and students are frequently persecuted, arrested, or tortured for their academic work, research, and publications; in others, the threats to academic freedom are more subtle, often driven by market dynamics and the increase of a corporate governance model of the university. (Furstenberg et al., 2020: para. 1)

Globally, academic freedom is under threat, with 3.6 billion people living in countries where academic freedom is completely restricted. Global levels of academic freedom increased to a peak in 2006, but by 2024 had declined to their 1973 level. In 2023, only 14 per cent of countries were classed as fully free, and 46 per cent were designated as completely restricted (Kinzelbach et al., 2024).

The European Parliament's *Academic Freedom Monitor 2023* noted that "the state of academic freedom in the European Union is relatively high on average compared to other regions and stable over time" (EP, 2024: 6). There were, however, nine European countries with a below-average level of academic freedom (Austria, Malta, Romania, Croatia, Bulgaria, the Netherlands, Greece, Poland, and Hungary). In the United States, researchers have been charged with using unethical research methods, and results were censored when election integrity and the role of mis- and disinformation was examined. Claims that research favours certain political parties circulate in the news media and via ostensibly independent organisations, in some cases leading to congressional investigations and

undermining public trust in information that circulates in the public sphere (American Sunlight Project, 2024). For example, Stanford University's Internet Observatory received pressure for undertaking rapid response tracking of electoral information (DiResta, 2024; Menn & Nix, 2023; Newton & Schiffer, 2024). In 2024, the Department of Justice brought a case against the Georgia Institute of Technology in relation to its cybersecurity lab for its claimed refusal to comply with Department of Defense security protocols while carrying out research commissioned by them (Abdalla & Abdalla, 2021; see also Anderson, 2024; Mueller, 2024). A case filed by the then attorneys-general of Missouri and Louisiana, both of whom had challenged whether President Biden had won the 2020 election, led to a Supreme Court ruling in 2024 clarifying when the government is permitted to communicate with researchers and social media companies undertaking research on mis- and disinformation about elections and vaccines (Tollefson, 2024). Until recently in Latin America, many universities had retained their autonomy, despite authoritarian governments. However, in Brazil, researchers at the Federal University of Rio de Janeiro's NetLab were targeted by the government in an effort to decrease their autonomy in setting their own research agenda on disinformation (Medronho, 2024). Big Tech companies' charitable giving to universities also can be seen as influencing research priorities, seeding doubt in research findings (Graham, 2024).

Retaliation against employees who become whistleblowers at tech companies such as OpenAI is well-documented – for example, Timnit Gebru, whose research on AI ethics and facial recognition at Google was suppressed, forcing her to leave the company (Hao, 2020; Knight, 2024). Evidence also suggests that the AI systems research field operates as an “economy of virtue” where “reputations are traded and ethical practices are produced in line with commercial decision-making” (Phan et al., 2022: 130; Eastwood, 2024). Proprietary AI systems make it difficult for researchers to interpret AI models or to devise public interest alternatives (Ahmed et al., 2023), though it is notable that a significant number of open-source LLMs are now available as options to proprietary LLMs, such as OpenAI's GPT-5 series.

Data governance, data justice, and democratic deliberation

The burdens of corporate data aggregation in today's data economy are disproportionately borne by those who are, or historically have been, subject to forms of social, economic, political, and cultural inequality or oppression. The government policymaking processes aimed at tackling these problems tend to exclude these same groups (Glasberg & Shannon, 2010; Glimmerveen et al., 2022; OHCHR, 2014). Even in local, national, or regional contexts, where laws and practices are designed to promote participatory politics and

democratic self-governance, policymaking processes are typically top-down and controlled by powerful elites, even when they are intended to perform as “representative” government (Glimmerveen et al., 2022; Wike et al., 2024). Those who are most vulnerable to the potential exposure of their information and to injustices and inequalities that come with massive data aggregation are least well-positioned to seek and obtain remedies for individual harms or to participate in civil society advocacy on these issues (Broomfield & Reutter, 2022; Eubanks, 2018; Georgiou, 2023; O’Neil, 2016; Ross Arguedas & Simon, 2023; Trappel, 2019). The space available for political communities to push for legislative reforms that might limit or shift data practices in fundamental ways is diminished as the private sector and public bureaucracies both become more dependent on digital infrastructures and algorithmic products (Mager & Katzenbach, 2021; Papaevangelou, 2023; Whittaker, 2021).

Corporate power and lobbying

Political processes involving government institutions relating to data governance are heavily influenced by corporate interests. This takes the form of lobbying, whereby powerful corporations engage highly paid professional lobbyists to meet with government officials and others who are well positioned to shape law-making (Bannerman et al., 2024; Popiel, 2018; Rankin, 2023; Ruohonen, 2003; Tarrant & Cowen, 2022). The largest tech companies spent close to 70 million US dollars lobbying in both 2022 and 2023, much of it to influence United States federal policy. These Big Tech companies also invest in lobbying to limit the scope of state-level legislation, as in the case of privacy and data-protection legislation, which has seen many years of congressional gridlock. Investigative reporters and civil society organisations uncovered “a coordinated, nationwide campaign by Big Tech” to shape state-level privacy laws (Feathers & Ng, 2022: para. 11), and of 14 state-level privacy laws, “all but California’s closely follow a model that was initially drafted by industry giants such as Amazon” (Fitzgerald et al., 2024: 6). Companies have a long track record of exploiting public emergencies to engage in turbo-charged lobbying that evades the democratic process (Klein, 2008). During the Covid-19 pandemic, Eric Schmidt, former CEO of Google, leveraged the panic of citizens and policymakers to campaign for massive public research and development expenditure and for the creation of dozens of public–private partnerships to embed data-intensive, corporate-owned tech platforms across multiple sectors – notably healthcare and education (Klein, 2020).

The lobbying arms of global tech firms also spend significantly to influence policy at the national level in countries outside the United States. An example is the legislative process that resulted in the passage of the European Union’s AI Act. In the lead-up to the adoption of the Act, European AI startups, such as Aleph Alpha and Mistral, as well as American Big Tech giants

(Google, Microsoft, and OpenAI), lobbied aggressively for amendments to draft legislation that would favour their own products and practices (Corporate Europe Observatory, 2024b; Perrigo, 2023). European watchdog organisations reported that technology companies had privileged access to European decision-makers (Corporate Europe Observatory, 2024a). For example, OpenAI lobbied officials to ensure that general purpose LLMs, such as ChatGPT, would not be treated as high-risk by default under the legislation (Perrigo, 2023). Lobbying is also strong in countries where emerging data markets promise to be very large. In India, Meta, Google, and Amazon lobbied aggressively against data localisation provisions in the country's data-protection bill (Business Standard, 2018; Sherman, 2022). When Brazil introduced legislation to combat "fake news" in 2023, United States-based tech companies campaigned against the bill. Google used its search engine home page to promote articles criticising the legislation and urging Brazilians to act against it (Boadle, 2023; Harris, 2023).

Technology companies also seek to influence legal frameworks that have an impact on profit margins by lobbying policymakers on issues related to global trade. For example, during negotiations around the Indo-Pacific Economic Framework, Big Tech companies pushed for specialised digital trade rules that would have limited the ability of 14 countries to enact their own regulations intended to restrain the activities of Big Tech companies at the national level (Birnbaum & Martin, 2023; James, 2022; Lawder, 2023).

These companies wield influence over policymaking processes through their entrenched relationships with policymakers. This takes the form of a revolving door between government and industry, whereby people leave high-level government jobs for high-paying corporate jobs, and vice versa (Alfonsi, 2019; Popiel, 2018). These relationships are the context in which informal lobbying through socialising takes place, which can affect legislative and regulatory outcomes (Li, 2023). The porous boundary between government and industry also results in more pervasive and difficult-to-document forms of influence, such as "corporate influence on regulators' systems of belief, policy preferences and ideological biases" (Popiel, 2018: 568; see also Caplan, 2023; Pickard, 2014; Teachout & Khan, 2014). This kind of influence can take years to manifest in concrete policy and is rarely discernible in any single legislative process.

The corporate grip on policymaking is tightened when companies have a monopoly on services that governments require, as in the case of the growing dependency of national security infrastructures on surveillance technology developed and maintained by major tech companies such as Palantir (Ball & Snider, 2013; Iliadis & Acker, 2022). As whole sectors within social, political, and economic systems become dependent on the data-intensive products and infrastructures developed and sold by corporations (for which they claim both the rights of intellectual property and the protection of trade secrets), the limits that policymakers are willing to impose on corporate data

practices narrow (Singh & Gurumurthy, 2021). These developments extend from services to infrastructure, as in the case of undersea cables that enable the United States to surveil other countries and facilitate countries such as China and Russia to do the same. As struggles over these “underground empires” ramp up, they become politically charged (Farrell & Newman, 2023).

Much attention focuses on American- or Chinese-owned Big Tech companies, but regional companies are also associated with dependencies created by extractive data economy models. An example is Mercado Libre, an Argentinian platform company (the fourth largest Latin American company) that operates in electronic commerce, fintech payment, and credit assessment markets. It uses Amazon and Google cloud services, but it also develops in-house data services that exploit users in the Latin American region (Franco et al., 2024). This highlights the multilayered and global nature of data dependencies and their complexities.

Data dependencies are created in other ways that impact how data is collected, stored, shared, and used. The procurement processes through which government agencies acquire or license corporate data-intensive products are incorporated into public systems (Calo & Citron, 2021; Crump, 2016; Faife, 2022). Procurement processes can be opaque, and even when transparent, they do not usually include a mechanism for public comment or participation. Vast digital bureaucracies have been put in place through procurement processes without the knowledge of those whose lives are most affected by them, and often without even the knowledge of elected officials who are supposed to represent constituent interests. The resulting digital bureaucracies profoundly change the way people relate to their governments, in some cases imposing substantial hurdles on those trying to access services and benefits to which they are entitled (Crump, 2016; Faife, 2022; Hardy & Williams, 2008). Treating data as an economic good that generates revenue and profits is hardwired into today’s information ecosystems. It results in endless amounts of monetisation as data becomes a pivotal asset in the data economy (Luchs, 2023), despite the fact that this is widely criticised for failing to deliver other societal goals (Purtova & van Maanen, 2024).

Towards data justice

The varieties of corporate data monopolisation (controlling capital and access to knowledge) mean that the Big Tech companies that produce and use data for directing commercial behaviour are complicit in maintaining an impoverished public sphere for discovering and sharing information (Franks, 2021). Other companies are capturing and cordoning off from public access rich sets of data that contain usable insights. This makes it extremely difficult for those inhabiting a digital space to enjoy experiences of shared inquiry:

In a world where scholarly research is paywalled, it's free to hop on YouTube to watch white supremacists spread racist theories about IQ and race. But, to read a scholarly article refuting the racist YouTubers' baseless claims with well-researched facts, the charge is USD 37.50 to overcome Sage Publishing's paywall. (Lamdan, 2022: 74)

The result is a radical difference between how information ecosystems are experienced when they operate within governance frameworks of institutional privilege and when they operate outside such frameworks. Where data is accessible through data repositories and networks designed to help easily use this data to extract information to build knowledge, people can see themselves as readers, thinkers, and participants in the public sphere. When most available information is a byproduct of corporate data practices that primarily aim to satisfy goals unrelated to the social project of knowledge production, people are forced to fight a system designed to treat knowledge-seeking as a consumer activity (Couldry & Mejias, 2019; Magalhães & Couldry, 2021; Schoon et al., 2020; West, 2019). Open access to information provisions do not resolve all the constraints on access to online reliable information, but they may at least reduce the cost barrier where a digital infrastructure is in place.

The comprehensive data surveillance that underlies digital infrastructures, and the way that companies aim to monopolise surveillance data to control who has access to what kinds of information, means that the role of data in information ecosystems should not be considered solely in terms of individual impacts. The inadequacy of conceptualising the harms of the data economy in individualistic terms is easy to see when thinking about the way that monopolistic corporate behaviour makes access to high-quality information contingent on social and economic privilege. However, a more multidimensional account of data harms is necessary to make sense of the asymmetrical structure of the markets the Big Tech companies are using data to create. The data collection practices of the most powerful technology companies are aimed primarily at deriving (and producing) population-level insights regarding how data subjects relate to others, not individual insights specific to the data subject. These insights can then be applied to all individuals (not just the data subject) who share these population features (Viljoen, 2021: 577). A primary motivation for companies to produce and aggregate massive amounts of data is to make classification predictions of characteristics and behaviours that facilitate targeting for economic or political purposes.

The most prominent approaches to data governance provide frameworks for protecting security (individual and/or state), property, and dignity and autonomy. Robust enforcement might improve outcomes for some individuals and communities, but these frameworks fail to provide a means for contesting datafication itself (Viljoen, 2021). While some existing or

proposed data governance frameworks address problems of improper economic or political influence, they do not challenge the underlying data practices that make such targeting possible. Nor do they take account of the way data practices that strive to shape individual behaviour according to AI tools' classification predictions can produce or entrench social injustices (such as wealth disparities or racial oppression) (Viljoen, 2021). Thus, population-level data injustice is related to, but conceptually distinct from, the individual harms that people suffer as a result of, for example, facial recognition algorithms trained on data that reflects racial bias (Mayson, 2018). The much-neglected problem at the population level is that the burdens of data production are borne disproportionately by certain groups. The case of a company called Fog Data Science illustrates the problem of population-level data injustice.



Fog Data Science, based in the United States, was founded by former Department of Homeland Security officials. Its digital tracking programme called Fog Reveal is sold primarily to law enforcement agencies to provide access to a database containing billions of records from 250 million mobile devices and enable them to conduct a variety of searches (including search by device ID). These searches can be used to develop "pattern of life analysis" – profiles of individual habits based on long-term behavioural data. Fog Data Science buys domestic and international location data from data brokers, which originates from over 700 smartphone apps that use ad ID. Ad ID is a way for advertisers to personalise offers for mobile device users using a random string of numbers and letters that attaches to the data that smartphone apps generate about users (year of birth, gender, search terms used, and location). Most mobile device users do not know about ad IDs – that apps are recording data about their movements or that this data is being purchased by data brokers and sold to the police. Those who are aware may think of it as a trade-off of their privacy for convenience. They cannot predict downstream harmful uses of their data. The use of ad ID non-transparently for policing risks disproportionately impacting some groups more than others. In the United States, a middle-aged white woman is much less likely to be targeted by police using Fog Reveal than a young Black man, for example, so Black men will disproportionately experience the harms not only of Fog Reveal, but of all the companies in the chain through which Fog Reveal obtains data (Cyphers, 2022; Greenberg, 2022; Turow et al., 2023).



Levels of awareness of, and tolerance for, privacy invasion set limits for data practices with drastically differential real-world consequences for various groups. Democratic data governance requires political structures that make it possible for communities to grapple with how data production and information extraction impacts the distribution of power between Big Tech companies or government agencies engaged in these practices and those who are already differently situated socially, culturally, and economically:

The status quo of data governance law, as well as prominent proposals for its reform [...] attempt to reduce legal interests in information to individualist claims subject to individualist remedies, which are structurally incapable of representing the interests and effects of data production's population-level aims. This in turn allows significant forms of social informational harm to go unrepresented and unaddressed in how the law governs data collection, processing, and use. (Viljoen, 2021: 578)

Given the global reach of Big Tech companies and the ease with which data is created, shared, transferred, and copied (Quintais et al., 2023), data governance interventions are essential that consider how the population-level effects of data practices manifest within, but also across, political boundaries. These interventions must also treat people's responses to corporate datafication practices as driven both by Big Tech profit motives and by a process whereby "data disaffection" and disengagement becomes part of a "structure of feeling" in a cultural dynamic, the perceived inevitability of which needs to be challenged (Grover et al., 2025).

Conclusion

This chapter examined how digital data, data-dependent digital infrastructures, data markets, and companies in the data business produce inequalities and how these are played out through contestations within information ecosystems. We emphasised the harms to democracy when these infrastructures become pervasive and increasingly invisible and difficult to contest. By focusing on corporate data monopolisation strategies, we explained why these strategies result in data dependencies and how they facilitate the development of biased AI systems as well as information markets that foreclose access to diverse sources of information, threaten research independence, and change the rules and norms concerning what counts as verifiable knowledge. We emphasised how Big Tech business models lead to the dependence of individuals, groups, and industry sectors on the products and services of these companies, showing how many current data governance practices preempt meaningful political deliberation.

The analysis in this chapter highlighted the significant risks to democracy when data and information are structured in ways that few understand or

have control over. This affects people's abilities to resist manipulations and to think and deliberate with others about the common good; it normalises data governance that is inconsistent with data justice. We emphasised the need for democratic deliberation, supported by multidimensional research that can inform efforts to devise and implement data governance that helps to alleviate population-level, as well as individual-level, discriminations and exclusions that can be attributed to the design of information infrastructures and to the way they are governed. We examine approaches to information ecosystem governance by states and Big Tech companies in the next chapter and frameworks for combating mis- and disinformation and data extractivism in Chapter 6.

Governing information ecosystems

Introduction

The term “governance” is used in multiple ways and is understood here to broadly encompass patterns of rules that underpin social orders (Puppis et al., 2024). This chapter emphasises the normative goals and rules that are embodied in approaches to governing information ecosystems with a focus on news media, AI systems, and data production and circulation. The literature in this field is extensive, and our analysis highlights selected prominent governance approaches including rules initiated by states, both through corporate self-regulation and through co-regulation involving governments and the private sector.

The emphasis in this chapter is on how governance approaches aim to cultivate “systems, institutions, and norms that enable quality and useful information to flourish” (Radsch, 2023a: para. 4) and on how they aspire to mitigate risks and harms associated with mis- and disinformation. Particular attention is given to efforts to balance the respective interests of governments, Big Tech companies, and news media organisations with the interests of individuals and groups within society. We also examine the extent to which human rights obligations are embraced. The effectiveness of these approaches in stemming the spread of mis- and disinformation is the subject of Chapter 6.

Varieties of information ecosystems governance

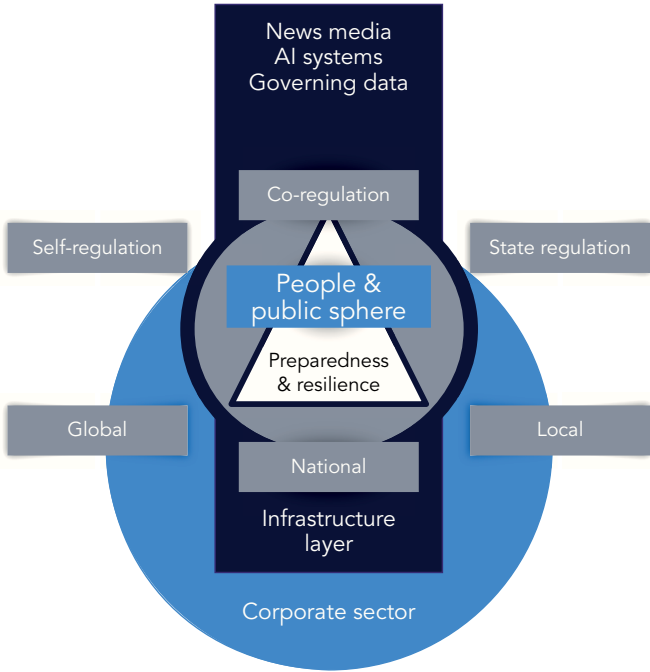
Information ecosystems are governed by numerous principles and rules. The companies whose platforms host and circulate content and other services and invest in and operate the underlying digital infrastructure, including the Internet, set rules and standards consistent with their interests in profit. Industry self-regulation comes with an expectation that the principles that companies adhere to, and their practices, should be consistent with internationally, regionally, or nationally agreed social, cultural, and political norms. Most Big Tech companies voluntarily employ some form of content governance and claim that this embodies human rights, including freedom of expression and data and privacy protection. Typically, they do not have

dedicated policies specifically for governing mis- and disinformation. The monopolistic practices of these companies, as discussed in Chapter 4, mean that self-regulatory practice often clashes with public interest (Kokshagina et al., 2023). For example, industry self-regulation may be intended to protect the integrity of elections but be limited due to its interference with business interests (i.e., in maximising and monetising user online engagement) or because of corporate indifference to the public interest (De Blasio & Selva, 2021).

States are primarily responsible for acting to protect human rights and fundamental freedoms, including in the digital environment. They have a *negative* obligation not to violate rights – including the rights of digital companies. They also have a *positive* obligation to protect human rights and implement them in practice. Governments may enact legislation that grants them direct authority to decide how information ecosystems should be structured and managed. Legislation is also introduced that must be adhered to by companies that operate the service applications and infrastructure layers of information ecosystems. These co-regulatory arrangements involve the state, the private sector, and often some form of civil society participation (De Blasio & Selva, 2021; Pickard, 2020b; Tenove, 2020). This governance approach aims to correct power asymmetries between the owners of digital platforms and other actors in the information ecosystem by establishing norms and rules for their behaviour.

Governance is a balancing act, and direct state regulation on its own can result in rights-infringing measures and partisanship. Co-regulation, in contrast, is often considered more flexible and inclusive because of its potential to leave less room for abuse and discretionary state measures (Dittrich, 2019; Durach et al., 2020). Co-regulatory approaches have become common in recent decades, although self-regulation by Big Tech companies is still the preferred approach in some countries (De Blasio & Selva, 2021). In addition to proactive approaches, governance also includes reactive measures to tackle the monopolistic power of Big Tech companies. Competition (or anti-trust) legislation may be used to curtail the monopolistic power of Big Tech companies in the expectation that the competitive supply of digital infrastructure and services will be more likely to create incentives for Big Tech companies to innovate and to operate in the public interest.

Figure 5.1 Simplified view of the governance of information ecosystems



In Figure 5.1, people and the public sphere are located within information ecosystems where global, national, and local governance – industry self-regulation, co-regulation, and direct state intervention – establish the principles and rules for how the providers of news media, AI systems, and data-dependent services operate. These rules influence the preparedness of societies and individuals to respond to the risks and harms associated with mis- and disinformation and their resilience and ability to participate in democratic processes (for background, see Brown & Marsden, 2023; Catalán-Matamoros, 2023; Flew, 2021; Flew & Martin, 2022; MacCarthy, 2023; McGuigan, 2023; Montero & Finger, 2021; Moore & Tambini, 2018, 2021; Tan, 2020; Tyagi et al., 2024; West & McAllister, 2023).

Governing principles and legislation

Governance principles for information ecosystems are established by institutions with global reach. The United Nations has set out guiding principles in its *Pact for the Future*, which includes a *Global Digital Compact*. The Pact calls upon all actors to embrace “full respect for international law”, “the pursuit and enjoyment of human rights and fundamental freedoms for all”, and “the responsible and ethical use of science, technology of innovation, guided by the principles of equity and solidarity” (UN, 2024b: 57–58). As

contests among stakeholders – public and private, individual and collective – over the role of Big Tech companies in information ecosystems have become prominent, objectives for information ecosystems have been established: 1) close all digital divides and accelerate progress across the Sustainable Development Goals; 2) expand inclusion in and benefits from the digital economy for all; 3) foster an inclusive, open, safe, and secure digital space that respects, protects, and promotes human rights; 4) advance responsible, equitable, and interoperable data governance approaches; and 5) enhance international governance of artificial intelligence for the benefit of humanity (UN, 2024b: 40–41).

Specifically concerning mis- and disinformation, a United Nations General Assembly comment emphasises that “*simple solutions* to this complex problem are likely to censor legitimate speech that is protected under international human rights law [emphasis added]” (UN, 2022: 11). Recognising this, a voluntary code of conduct on the integrity of information states that “all stakeholders should refrain from using, supporting or amplifying disinformation and hate speech for any purpose, including to pursue political, military or other strategic goals, incite violence, undermine democratic processes or target civilian populations, vulnerable groups, communities or individuals” (UN, 2023a: 21) and it sets out principles (see Table 5.1).

Table 5.1 United Nations principles for information integrity

- Commitment to information integrity
- Respect for human rights
- Support for independent media
- Increased transparency
- User empowerment
- Strengthened research and data access
- Scaled-up responses
- Stronger disincentives
- Enhanced trust and safety

SOURCE: UN, 2023a: 21

The *Global Digital Compact* calls upon United Nations member states to “work together to promote information integrity, tolerance and respect in the digital space, as well as to protect the integrity of democratic processes” (UN, 2024b: 49). Digital platform companies are called upon to work towards ensuring that their uses of technology do not amplify mis- and disinformation and other harms, and it is acknowledged that this will require risk mitigation and redress measures (see also Chapter 1).

Similarly, the OECD's Council recommendation on information integrity addresses the need to enhance the transparency, accountability, and plurality of information sources, foster societal resilience, upgrade governance measures, and encourage appropriate institutional arrangements (OECD, 2024b). The OECD (2024a: 17) also emphasises that,

what makes content-specific regulatory responses particularly complex is not only that defining what content may be restricted without infringing upon freedom of expression is difficult, but also that illiberal regimes can co-opt laws to combat disinformation developed in countries with effective checks and balances to legitimise their own antidemocratic practices.

Bilateral initiatives aim to protect democracies from mis- and disinformation campaigns of foreign governments (Wintour, 2024). All these statements of priorities involve voluntary commitments which generally have less traction than rules and norms introduced nationally or regionally.

Governing network infrastructure

On the infrastructure layer of information ecosystems, Internet connectivity and access to information are central to how information ecosystems are experienced. Network neutrality policies and regulations and “zero rating” data packages typically involve co-regulation but may also involve states acting directly under legislation permitting them to establish rules for Internet access and use. These policies condition whether and how people can generate and amplify the circulation of all kinds of information.

Network neutrality is the principle that Internet service providers (ISPs) should treat all data (information) that flows through their networks without discrimination to maintain an open Internet. This principle, which initially emerged in the Global North and was intended to ensure equality of access to information, is controversial, because it can make it more challenging for ISPs to differentiate their prices for data usage for information provided by different providers. For example, zero-rating pricing plans have been introduced in many countries in the Global Majority World. “Zero rating” refers to ISPs offering access only to certain services and placing data usage caps on accessing those services. Data usage pricing policies are used to limit or discourage access to certain kinds of information, that is, these policies discriminate in ways that are considered incompatible with an open Internet. Combined with algorithmic personalisation tools, people can be locked into “walled gardens”, and it is more costly for them to access diverse sources of news that are not part of the package permitted by a zero-rating data service contract. Online users may then restrict themselves to information that is part of their pricing package to minimise costs (Palmer & Toff, 2022; Aharoni

et al., 2021). There is debate about whether network neutrality policies and zero-rating result in exclusion due to unaffordable network access and whether access to information is unjustly reduced. These policies do mean that high-choice news media environments are not available to everyone, even when the Internet is available and reasonably affordable (see Chapter 1; on network neutrality, see Baranes, 2014; Bauer & Knieps, 2018; Economides & Hermalin, 2012; Hildebrandt & Wiewiorra, 2024; Lee & Shin, 2016; Marsden & Brown, 2023; Menon, 2021; Pickard & Berman, 2019; Winseck & Pooley, 2017; Wu, 2003; Yoo, 2024; on zero-rating, see Gerpott, 2018; Hoskins, 2024; Jaunaux & Lebourges, 2019; Jordan, 2017; Krämer & Peitz, 2018; Marsden, 2016; Mattelart, 2023).

In India, the country's diverse socioeconomic context means that Internet access is treated as a critical developmental tool and an effort has been made to ensure equal access by banning a two-tiered Internet. This has required a difficult balance between equity considerations and innovation in service provision.



Network neutrality and zero-rating issues in India have been intensely scrutinised by online content companies and ISPs. The Telecom Regulatory Authority of India (TRAI) consulted on network neutrality in 2016 and 2017 after a #SaveTheInternet campaign was launched against Facebook's Free Basics proposal to provide low-cost, subsidised access to a few selected services for lower-income countries in Asia and Africa. Its zero-rating service would have given preferential treatment to certain services, and, in India, this raised concerns about fairness and competition. Network neutrality policy debate has focused on traffic management, that is, the prioritisation of certain types of Internet traffic over others, which potentially disadvantages some users or competing services. Conflicts have arisen between proponents of an open Internet, where all data is treated so that it flows equally through the network, and ISPs claiming the need to manage network traffic congestion and block illegal content. TRAI banned discriminatory tariffs in 2016, effectively prohibiting zero-rating offerings like Free Basics, but the Cellular Operators Association of India argues that the policy is limiting the introduction of lower-cost access that could help to bridge the digital divide (Eisenach, 2015; Menon, 2021; Mukerjee, 2016; Prasad, 2018).



Network neutrality policy and zero-rating are widely discussed in other countries and regions (e.g., South Africa, South Korea, and Latin America), with disputes over the public value of the Internet and with digital activism

directed to resisting Meta's Free Basics service (Nothias, 2020; Robb & Hawthorne, 2019; Shahin, 2019; Shin & Lee, 2017). In jurisdictions that do allow zero-rating, regulatory bodies, such as the Independent Communications Authority of South Africa, have provided guidelines to align these practices with public interest goals such as enhancing education and public health access to information or to prevent anti-competitive behaviour.

Techniques for closing off Internet access, whether via zero-rating, throttling traffic, or other means of fragmenting the Internet, are widely in use (Boas, 2006; Howard & Hussain, 2013; Kalathil & Boas, 2003; Shahbaz et al., 2022). There are examples of Internet shutdowns and social media blocking during elections with political unrest and protests in countries as diverse as Belarus, Iran, Myanmar, Turkey, Vietnam, and Zimbabwe (Akser & Baybars, 2023; Mare, 2020; Sinpeng, 2020; Rying et al., 2022). Authoritarian regimes restrict access to information, as in the case of China's Great Firewall, by heavily regulating national Internets, using content governance to favour or censor political speech, and for surveillance (Keremoğlu & Weidmann, 2020). Restrictions on information access include Iran's initiative to create a national information network by requiring service providers to locate servers inside the country and by increasing the cost of global Internet traffic (Motamedi, 2024).

In the United States in 2015, the Federal Communications Commission established network neutrality rules to prevent ISPs from discriminating to favour their preferred online service providers. When network neutrality rules have been in place, a majority of Americans report that they either understand or support the policy, even though its enforcement has been inconsistent and has impacted the quality of service they experience (Vogels & Anderson, 2019; see also Program for Public Consultation, 2022). The Federal Communications Commission's network neutrality decision was repealed in 2017, reinstated in 2024, and then blocked by a federal court. There is some evidence that without network neutrality rules, ISP data throttling and prioritisation can disadvantage smaller content providers and reduce consumer choice (Greenstein et al., 2016). In the European Union, the Open Internet Access Regulation provides rules on network neutrality, and the Body of European Regulators for Electronic Communications sets guidelines requiring that ISPs should not favour specific service providers, and any zero-rating pricing should not undermine net neutrality (BEREC, 2024a, b). This policy is controversial, with critics arguing that the policy allows for subtle discriminatory data traffic prioritisation, although there is a high compliance rate.

These approaches to governing the infrastructure layer highlight the need for adaptable and context-sensitive regulation so that the benefits of digital inclusion and access to diverse sources of information are balanced against the risks of market monopolisation. One strategy is to require ISPs to publicly disclose their data management practices and zero-rating agreements. Another is to include civil society and industry voices in the policymaking process.

Decisions about network neutrality and efforts to limit how it operates are influenced by deliberations on digital sovereignty with outcomes that differ depending on a country or region’s political and economic context (Afina et al., 2024; Bria et al., 2025; Kokas, 2022; Kumar & Thussu, 2023; Stefanija & Pierson, 2023).

Governing digital platforms

UNESCO has developed Guidelines for the Governance of Digital Platforms, which establish voluntary principles and duties, responsibilities, and roles for stakeholders, with the aim of safeguarding freedom of expression, access to information, and other basic human rights (see Table 5.2). It has also produced guidance for policymakers to address hate speech through educational initiatives (UNESCO, 2023a, b).

Table 5.2 UNESCO principles for governing digital platforms

- Platform owners conduct human rights due diligence
- Platform owners adhere to international human rights standards, including in platform design, content moderation, and content curation
- Platform operations are transparent
- Platform companies make information accessible
- Platform owners are accountable to relevant stakeholders

SOURCE: Adapted from UNESCO, 2023b: 20

Many countries seek to establish rules to govern how information ecosystems develop by moderating the behaviour of the Big Tech companies when it is inconsistent with human rights standards or is anti-competitive (Krotoszynski et al., 2025).

The European Union’s regulatory framework includes the Digital Services Act (DSA), the Digital Markets Act (DMA), the Data Governance Act, the Data Act (EC, 2022a, c, d, 2023), and other legislation sharing a foundation in aiming to safeguard citizen rights, and this framework shapes approaches to governing content at the national level, as in the case of, for example, the German Network Enforcement Act of 2017 (Government of Germany, 2017), the French Organic Law No. 2018–1201 (Government of France, 2018), and Hungarian legislation (Stolton & Makszimov, 2020) (see also Akman, 2022; Botta, 2021; Broughton Micova & Jacques, 2020; Church & Pehlivan, 2023; Galantino, 2023; Just, 2022; Mansell, 2021; Moreno Belloso & Petit, 2023; Nenadić et al., 2023).

The DSA is “a horizontal framework for regulatory oversight, accountability and transparency of platforms and search engines” (Nenadić et al., 2023: 8; see also Turillazzi et al., 2023), and its provisions govern the algorithms used

in automated content moderation, with binding obligations to remove illegal content, safeguards to respect freedom of expression, and substantial penalties for failure to comply. The DSA (EC, 2022c) governs the interactions between users and platforms – including the degree to which algorithmic recommender systems are used. Article 14 mandates that platforms must consider user interests in content moderation and complaint handling, referencing fundamental rights, such as freedom of expression. For example, very large online platforms (VLOPs) – defined as platforms with more than 45 million users per month – are required to respect fundamental rights based on comprehensive risk analyses and mitigation strategies (Art. 34.1) (Nenadić et al., 2023; Mahler, 2022). VLOPs and search engines are expected to address the systemic risk of “negative effects on civic discourse and electoral processes” associated with their services (EC, 2022c: Art. 34.1c). They must adhere to a benchmark for processing valid notifications for removal of illegal hate speech in less than 24 hours. If a platform considers that content is not compliant with its terms and conditions, it may proceed with deletion or restriction. The aim is to create a safer digital space within which the fundamental rights of all users of digital services are protected (EC, 2022c; Reyna, 2024; Broeders et al., 2023: 1272; Mansell, 2021).

The DMA is addressed to the monopolistic behaviour of the largest digital platforms, with the aim of establishing a level and competitive playing field, that is, contestable markets, by constraining the practices of Big Tech companies with gatekeeping power over core platform services. The goal is to promote “innovation, high quality of digital products and services, fair and competitive prices, as well as high quality and choice for end users in the digital sector” (EC, 2022a: para. 106) by addressing imbalances in bargaining power and unfair practices so that greater choice is available to platform users. There are sanctions against self-preferencing, the largest gatekeepers must enable the interoperability of services, and other measures aim at achieving a balance between business and individual (or collective) interests (Brown & Marsden, 2023; Crémer et al., 2019).

In addition, the Treaty on the Functioning of the European Union (TFEU) applies to the conduct of gatekeeper companies (EC, 2012: Art. 101, 102). The scope tends to be limited to cases where the dominance of specific markets can be evidenced through investigation. The criteria for establishing market dominance are slowly being modified to reflect the complexities of the business models of the Big Tech companies and the fact that users of many online services do not pay for this use. For example, in Germany, non-price issues, such as access to data, have been treated as a potential criterion for determining market power, and member states are introducing modifications to enable them to bring actions against the owners of digital platforms more easily (Just, 2018, 2022). Competition policy provides tools that can be used to mitigate harms arising from market power, and it may involve remedies, including corporate divestment, fines, or behavioural requirements. Cases have been

brought against Google for search and advertising practices (resulting in a 2.95 billion Euro fine in 2025 for favouring its own online advertising services), Google and Apple for app store rules for participation, Meta for data collection and processing practices, and Amazon for its treatment of companies that use its online marketplace (Nicoli & Iosifidis, 2023).

In the United States, until recently, there have been few efforts to curtail the market power of dominant digital platform companies. Although the companies insist that they provide their customers with convenient ways to access digital content and buy goods online, monopolistic market power has allowed them to refine their business models to maximise user engagement and monetise data for profit, as discussed in Chapter 4. The Big Tech companies have acquired or suppressed competitors, favoured their own products and services, and downplayed or disavowed responsibility for harms linked to data collection and processing operations. More aggressive application of anti-trust law was encouraged under the Biden Administration, with cases brought against the platforms by the Department of Justice and the Federal Trade Commission in a bid to curtail monopolistic behaviour (FTC, 2024; see also Stigler Committee, 2019; Wu, 2018). These kinds of cases continue to be prosecuted under the Trump Administration.

When cases are brought that confirm the harmful exercise of monopoly power by the platforms, this has the potential for substantial impact. The courts in the United States assess whether a company's general search engine that crawls the web and generates its own results, (e.g., Google or Bing and Yahoo, which syndicates results) has exercised such power by examining different categories of search ads that generate revenue. Some are general search text ads, the most predominant form of sponsored ads, while others are shopping ads or product listings – and there are other types as well. An example is a court ruling in August 2024 that Google was a monopolist in the general search text ad market. It was not found to be a monopolist in the market segment that is based on the signals provided by users' online interactions and the company's algorithm-driven responses to these signals. The general search ad market excludes display ads, retargeted display ads, and non-search social media ads, that is, ads that rely on "indirect signals to decipher a users' latent intent" based on a user's past online interactions (US District Court, 2024: 168). In another case brought by the Department of Justice, a District Court found that Google had monopolised open-web digital advertising markets, which has enabled the company "to censor and even deplatform American voices" (US Department of Justice, 2025). Penalties might have included breaking up the company or other measures aimed at enforcing changes in its algorithms (Radsch, 2024). In the event, Google was not required to divest, but it was ordered to share some of its data with its competitors in a decision that was seen by critics as a light-touch outcome.

Proposals for new legislation in the United States are considered at the federal level to tackle Big Tech power. Few have been signed into law, but

they have called for prohibitions on large platforms that preference their own products, encouraged platform interoperability, and sought to restrict platform use of non-public data. In May 2025, with bipartisan support, the Take It Down Bill was signed into law (US Congress, 2025). It creates an obligation on request to take down unlawful and nonconsensual “deepfake” or “revenge porn” content online. Not without controversy, critics, including the Center for Democracy and Technology and the Electronic Frontier Foundation, argue that the use of vague language opens the door to platform deletion of legitimate images such as LGBTQ+ content or non-public content stored on company servers (Ortutay, 2025).

It is argued that the use of services provided by digital platforms provided by Big Tech companies has reached a point where they are essential public services that should be subject to the same regulations in the United States as public utilities that operate as “natural” monopolies (Pickard, 2022a). However, whether the Big Tech platforms meet the threshold for being classified as an essential service is disputed, with some arguing that treating them in this way could entrench their monopolistic positions.

None of these platform governance approaches radically changes an overarching commitment in the United States and in the European Union to promoting rapid innovation in digital technologies that use opaque algorithms for profit. This remains a powerful mobiliser of investment in future generations of technologies. This tension is also present in other regions and countries when an effort is made to balance human rights protections with economic interests.

Governing news media

The application of governance rules to the news media is expected to preserve and maximise diversity and a plurality of voices in the public sphere within the framework of internationally agreed upon rights and responsibilities. The news media in Western democracies have been largely self-governing to protect their independence, although news media freedoms historically have never been absolute, and the privileges and duties of the journalism profession vary around the world. Responses to changes in the relations between news organisations, platforms, and state actors that are engaged in producing and circulating news need to be guided by human rights standards and normative expectations, although there are deviations in practice (Tambini, 2021a).

In some countries and regions, action is being taken to promote news media pluralism and media freedom and to counter mis- and disinformation (Paal, 2017). The European Union introduced a Democracy Action Plan with measures to promote free and fair elections, strengthen media freedom, and counter mis- and disinformation (EC, 2020b). The Media Freedom Act (EC, 2024b) aims to protect journalists’ work, secure the independence of public media, and increase the transparency of private media ownership. It

requires a fair allocation of state advertising revenue to news media producers and obliges member states to implement media concentration assessments (although it neither prevents media concentration nor sets a threshold). It also introduces a media pluralism test, requiring member states to examine media mergers in relation to media pluralism and editorial independence and requiring market competition assessments. Despite being a substantial shift away from the previous hands-off approach, the act excludes user-generated content unless it is uploaded for financial or other consideration, purely private correspondence and services that do not have “provision of programmes or press publications as their principal purpose” (EC, 2024b: para. 9), corporate communication, and informational or promotional materials, although it does cover media produced by freelancers (Brogi et al., 2023; Parcu et al., 2022). Broughton-Micova and Schnurr’s (2024) research emphasises the need for benchmarking and empirical evidence to track progress towards “good” electoral and information integrity as a result of both the DSA and the Media Freedom Act.

When news media are governed as a “public good”, the expectation is that this will help to maintain independent public service media organisations (UNESCO, 2022d). In countries and regions where public service media is reasonably shielded from political pressure, these organisations are a vital component of the information ecosystem (Michalis & D’Arma, 2024). In countries where governments put pressure on news media organisations, both privately owned media and public service media can fail to meet normative expectations (Farahat, 2021). The regulation of news media can backfire when it is a pretext to consolidate state power and control over information flows leading to censorship and repression or to more subtle forms of leverage that hold news media organisations in check. For example, in Cuba, the state maintains control over the media (also dominating artistic and intellectual affairs) by prohibiting private (legacy and online) media outlets under the 2019 Constitution, which classes them as funded by “enemies of the state” (Garcia Santamaria & Salojärvi, 2020; Romeu, 2023). In Hungary, efforts to control regulatory bodies and a concentrated media market have helped to consolidate domination by the ruling party (Polyák, 2019).

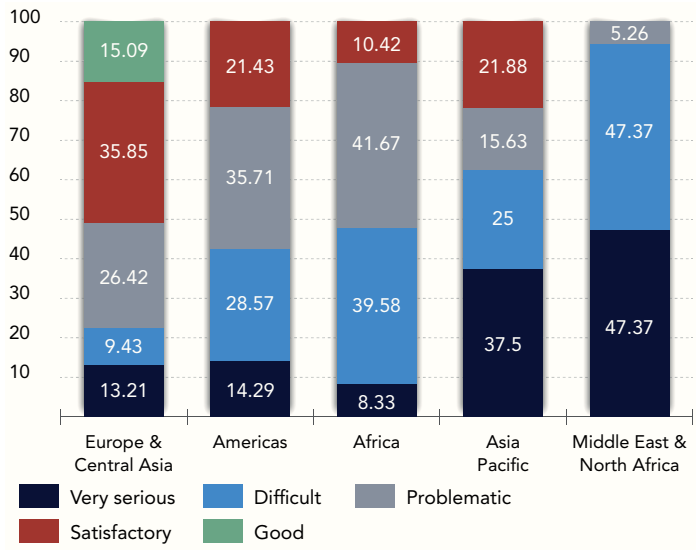
When Apartheid ended in 1994 in South Africa, media governance arrangements were introduced. The 1996 Constitution gave unprecedented levels of freedom to media organisations, emphasising the priority to build an ethical, independent, and publicly accountable news media, moving from media self-regulation to co-regulation (Wasserman, 2020b). Since then, there have been complaints that the news media serves the interests of an elite, that disadvantaged community voices are not represented, and that the public broadcaster, South African Broadcasting Corporation (SABC), lacks independence. SABC has been criticised for being too soft on the

elected government, with debate focusing on whether the news media should be more critical of the democratic government or protect democracy by supporting it (Wasserman, 2020b).

In Turkey, the government has fostered a favourable news media by leveraging structural, legislative, and illegal measures to benefit the ruling party. After a failed coup in 2016, a restructuring of the media system led to greater repression through measures including economic incentives, structural support for the media favouring the ruling party, and control of regulatory bodies (Akser & Baybars, 2023). In Venezuela, a legal reorganisation under the Hugo Chávez government (1999–2013) shifted the media system from private dominance (opposing the government) to state dominance (supporting the government), without alleviating the political and economic pressures on news media organisations. Under Nicolás Maduro (2013–2019), the news media experienced further government pressures (Gracia Santamaria & Salojärvi, 2020). In both Vietnam and Singapore, media regulations apply to govern censorship, ownership, and personnel management, and there are other repressive instruments. Vietnam’s approach is considered coercive, while in Singapore, political norms are enforced implicitly by embedding stakeholders with financial interests in the media system (Haenig & Ji, 2024). These repressive measures infringe on human rights and create a space in which to produce mis- and disinformation and promote its circulation through both legacy and online news media.

When news media regulation infringes on human rights, this contributes to democratic fragility, since verifiable news and informed public opinion are essential for a functioning democracy, which requires high-quality public debate and deliberation and accountable representation (Pickard, 2022a, b; Tambini, 2021b; Tenove, 2020). The degree of press freedom influences whether journalists feel secure and safe enough to produce content that fairly reports on country conditions. Reporters Without Borders defines press freedom as “the ability of journalists as individuals and collectives to select, produce and disseminate news in the public interest independent of political, economic, legal, and social interference and in the absence of threats to their physical and mental safety” (RSF, 2022: Aim section). Press freedom is measured using indicators of the political context, legal framework, economic and sociocultural context, and safety. Figure 5.2 shows the state of world press freedom in 2024, with a global average decline of 7.6 points on the political indicator since 2013. Only 15 per cent of countries achieved the (green) good press freedom status (RSF, 2024; see also RSF, 2025).

Figure 5.2 World Press Freedom Index, 2024 (per cent)



SOURCE: RSF, 2024; redesigned by Matthieu Mercier and adapted by Nordicom.

If news media governance is to uphold the right to freedom of expression and freedom from harassment and violent attacks, it must tackle structural inequalities, technological change, political alignments, and social transformation (Davis et al., 2020). Governance measures can help to strengthen newsroom practices, contribute to news organisation financial sustainability, prevent the silencing of journalists, and strengthen the integrity of information. The Organization for Security and Co-operation in Europe’s (OSCE) Group of Friends on the Safety of Journalists has called for the implementation of commitments to protect journalists (Kamp, 2023). The United Nations’ Plan of Action on the Safety of Journalists and the Issue of Impunity (UN, 2023c) and the Freedom Online Coalition (n.d.) also contribute to this area. These initiatives aim to acknowledge the subtleties, trade-offs, and negotiations that journalists engage in when “they try to understand, anticipate, leverage, or resist how social media platforms see and organize news” (Ananny, 2023: 1570). Protecting independent journalism and independent news organisations is essential, but this requires “*freedom, funding, and a future* [emphasis original]” (Nielsen et al., 2019: 11), which is not improving in many countries.

Without sufficient funding, news media “independence can be easily compromised and pluralism becomes a shadow of what it should be” (UNESCO, 2022d: 7). Intermediary companies may claim to free news media organisations from their financial dependence on the largest platforms. For example, content recommender platforms such as Taboola and Outbrain operate in Israel and embed their widgets in news media websites such as the BBC, CNN, or *The*

Washington Post, NBC News, and FOX News. They matchmake between advertisers, publishers, and data brokers. However, they also blur news with sponsored content, making it difficult for audiences to detect the difference (Ratner et al., 2023). Development aid has also helped to sustain journalism and independent media in the Global Majority World by supporting local initiatives, but from 2010 to 2019, media aid worldwide stagnated as a proportion of total aid (0.3% of 200 billion US dollars, an average of 300–400 million US dollars annually) (Myers & Gilberds, 2024). In early 2025, the Trump Administration froze 268 million US dollars in foreign aid, which had been supporting independent media outlets in more than 30 countries. By March, an 80 per cent cut in donations had been confirmed, placing independent media in jeopardy and closing many services (Daniels, 2025).

Governing AI systems

Regulators and policymakers face substantial challenges in defining rules for AI systems when they seek to balance the potential benefits to be gained from encouraging innovation against the risk of harm to individuals, businesses, and society when there is a lack of regulation to protect them. International bodies such as the Council of Europe, the OECD, and the United Nations and its agency UNESCO are active in defining principles and standards designed to protect human rights against the negative impacts of AI systems (Bello y Villarino, 2023). Professional societies, such as the Institute of Electrical and Electronics Engineers (IEEE), are also active, especially in relation to electronic warfare (Koene et al., 2018). The literature on AI systems governance is developing rapidly, with growing numbers of contributions from legal scholars (see Bullock et al., 2022; Crawford, 2021; Gritsenko et al., 2022; Gunkel, 2024; Paul et al., 2024; Quintavalla & Temperman, 2023; and for law scholarship, see Bullock et al., 2022; Custers & Fosch-Villaronga, 2022; De Bruyne & Vanleenhove, 2021; Księżak & Wojtczak, 2023).

It is estimated that some 800 AI policy initiatives in at least 60 countries have been discussed since 2016, with the aim of tailoring AI governance to country conditions in a way that respects human rights and results in transparency and accountability (Roberts et al., 2023). The normative goals of AI systems governance still need clarification, and there is a need for a shared understanding of the societally appropriate balance between economic and innovation imperatives and human rights (Cornils, 2020). The challenge is to decide on that balance and then design legislation that can deliver and sustain it against the risks, for example, of regulatory overreach by governments, and particularly by authoritarian governments (Marda & Milan, 2018).

Before the emergence of AI-related regulation, information ecosystems were already regulated at various levels: international law, regional standards, and national laws. Initiatives to regulate the impact of AI on societies have started to home in on transparency requirements, training data disclosures,

and risk assessment obligations, many of which aim to guard against the risks and harms of mis- and disinformation. Normative approaches on the international level include:

- the United Nations resolution, “Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development” (UN, 2024c),
- the OECD’s Recommendation of the Council on Artificial Intelligence (OECD, 2022c),
- UNESCO’s Recommendation on the Ethics of AI (UNESCO, 2022c),
- the G20 AI Principles (G20, 2019),
- the G7 Hiroshima Process (G7, 2023),
- the AI Safety Summit Bletchley Declaration (including principles for GenAI) (UK DSIT, 2023), and
- the Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law (Council of Europe, 2024).

Policymakers and their regulatory institutions face two connected challenges when developing rules for governing the impacts of AI on information ecosystems (De Gregorio, 2023). First, AI is not a static product that need only be regulated once, and regulation must focus on the evolution of AI systems during their whole lifecycle – that is, throughout the design, development, and deployment phases. Second, the impacts of the use of AI systems are considered more determinative of regulatory needs than abstract characteristics of a system (which are bound to change). For this reason, most regulatory approaches are risk-based and are used to define AI systems requirements based on the level of risk a system is judged to pose (Cole, 2024). These approaches are seen by some as “not primarily to manage risk but instead to ensure legislative proportionality” that will avoid stifling innovation (Mahler, 2022: 247).

In Europe, there are prohibitions on certain kinds of AI applications, stringent rules for high-risk systems (such as certification schemes) and transparency obligations that mirror concerns about the “black box” nature of advanced machine learning systems. The European Union’s AI Act, which entered into force in 2025, includes obligations on the part of AI developers to actively protect human rights (EC, 2024b, c; Larsen & Küspert, 2024). This Act classifies AI applications into several risk categories, with more extensive obligations for higher-risk applications.



AI system risk categories

In the first category, certain AI systems are deemed so risky as to be unacceptable: An AI-based assessment of individuals' behaviour by government agencies, that is, when they influence the "free will" of users or contain social scoring. Under the AI Act's scope, their use is prohibited in the European Union. The next category includes "high-risk" AI systems, which are listed in Annexes II and III of the Act. Annex II features a list of existing European Union regulations that require a conformity assessment for products that bear specific risks. If an AI component is part of these products or the product itself, it is considered a high-risk AI system. For the list in Annex III, the context of use is more relevant, that is, it is not the AI system itself that is considered risky, but the domain in which it is applied. Eight domains are named in which certain AI systems are high risk, such as those involved in decisions about access to education or employment. A particularly large number of applications that are considered high risk are those involved in law enforcement or migration. If an AI system falls into this category, manufacturers and users must adhere to compliance obligations, such as having risk governance and quality management systems in place and registering the AI system with the European Commission (Annexes I and III of the Act refer to harmonisation with European Union legislation and list high-risk AI systems requiring third-party conformity assessment). The third category includes "low-risk" AI systems, for which the Act requires only transparency obligations and thus, significantly fewer requirements than for those in the high-risk category. This means that providers of AI systems that 1) interact with humans, 2) are used for emotion or biometrics recognition, or 3) that generate deepfakes must notify their users that the content was generated by AI.



Not regulated by the AI Act are "risk-free" AI systems that include, for example, spam filters for e-mail clients. Here, the risk for users is considered so small that no regulation is envisaged. For emerging AI systems not previously addressed, the AI Act stipulates that they must be categorised as high-risk AI systems if they can negatively affect fundamental rights. This classification imposes substantial compliance duties on both providers and users of these AI systems. Article 13 requires providers of high-risk AI systems to transparently outline the risks these systems pose to fundamental rights when employed, and Article 14(2) mandates human oversight of high-risk AI systems to safeguard fundamental rights.

The AI Act also references fundamental rights. These often serve to clarify the broader context and rationale for specific provisions at the European Union level, highlighting the potential of AI systems to impact fundamental rights. For instance, recitals (legislative texts) address the risks of AI systems being used for manipulative or exploitative practices. One criticism voiced against regulating only primary uses of AI models, mainly exercised through quality assurance of their training data, is that risks of secondary use, where a model used in an AI system is applied in a way that its developers did not intend, may go undetected. A solution would be to focus on the concept of purpose limitation for AI models, which would leverage existing data-protection approaches (Mühlhoff & Ruschmeier, 2024). Some researchers are encouraging more stringent AI systems rules, arguing that lawmakers should learn from ineffective and missing regulations during the early days of social media when they failed to address the underlying business model that led platforms to prioritise the data-driven monetisation of user attention (Sanders & Schneider, 2024). There are calls for standards in this area, and some argue that companies will use voluntary standards to evade regulations on AI systems development (Han et al., 2022; Lewkowicz & Sarf, 2024; Schwartz et al., 2022).

Some critics are not satisfied that policy interventions such as the AI Act will be effective. They argue that regulations on the use of AI systems in the generation, moderation, and curation of content streams need to be better informed by evidence of the nature of the threats and effectiveness of mitigation measures (Nannini et al., 2024). Others point to the need for effective verification processes for assessing compliance without which “this regulation risks being as ineffectual as GDPR [the General Data Protection Regulation], yielding only marginal improvements in ethical AI system practices” (Schmitt et al., 2024: 36),

Differences are apparent in AI governance initiatives in the European Union and the United States. In Europe, the AI Act excludes AI systems developed “exclusively” for military or national security purposes, although the distinction between AI technologies for civil and military use is difficult to make, and dual-use technologies could potentially fall within the scope of the Act. In the United States, President Biden’s Order on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (Federal Register, 2023) emphasised “AI for good” and the need to mitigate its substantial risks through the development and use of AI safety measures. It also set out obligations for AI developers of dual-use foundation models to show that these would not lead to violations of federal laws on civil rights, discrimination, or security. A dual-use foundation model was defined as a model,

trained on broad data; generally uses self-supervision; contains at least tens of billions of parameters; is applicable across a wide range of contexts; and that exhibits, or could be easily modified to exhibit, high levels of

performance at tasks that pose a serious risk to security, national economic security, national public health or safety, or any combination of those matters. (Federal Register, 2023: para. k)

In January 2025, however, President Trump rescinded many features of the earlier Biden order, including those on the development of dual-use technologies, with the aim of removing regulatory barriers to United States leadership in AI markets, placing the focus on competition and deemphasising or cancelling oversight, risk mitigation, and “responsible AI” developments (The White House, 2025a).

AI systems, and LLMs specifically, raise a host of issues around intellectual property rights. Human rights-friendly copyright frameworks for GenAI are under consideration internationally (Geiger, 2024), with an emphasis on the rights of human creators. UNESCO has considered the impact of AI systems on cultural production (Kulesz, 2018), and the World Intellectual Property Organization has acknowledged the significant legal uncertainties and the need for governance to vary by jurisdiction (WIPO, 2024). In the European Union, if a work is created by AI, it is not subject to copyright, but there is scope for application of the law if a creator is deemed to have given explicit instructions to an AI system. The AI Act says that text and data-mining operations used to collect training of LLMs must receive consent unless they are subject to exemptions – which so far seem to apply – but companies must document their use of data. Numerous court proceedings are underway to clarify whether the Big Tech companies should be able to freely access copyrighted works without compensation to creators. In the United States, copyright protection is not offered to creations produced by GenAI, but it is not clear what liability OpenAI and other firms have now, or will have in the future, for scraping data to train LLMs. Legislation is being presented to Congress, but none has succeeded in becoming law at the time of writing, and there is persistent controversy about the interpretation given to “fair use” of online text and data.

Many other regions and countries are putting AI system strategies and governance frameworks in place (Findlay et al., 2023). Although AI systems are largely developed by companies in the Global North and China, organisations in African and Latin American countries and elsewhere in Asia are developing applications using large datasets and machine translation tools, and there are calls for the localised development of AI applications (Okolo, 2023). For the African Union, a Continental Artificial Intelligence Strategy has been agreed (African Union, 2024), although AI regulation in Africa faces challenges of “weak institutional frameworks, limited judicial capacity, lack of expertise from policymakers, fragmented laws, and poor enforcement mechanisms, where laws, even if existing, are seldom applied” (Tech Hive Advisory Center for Law & Innovation, 2024: 21). In China, it is argued that the emphasis of AI governance is on strong binding regulations

(Zhu, 2022). AI legislation in Brazil has followed the European Union's AI Act in adopting a risk-based approach with a list of prohibited applications. It differs, however, in guaranteeing individual rights accompanied by judicial and administrative mechanisms of enforcement. These include the right to contestation and human intervention, emphasising due process for people affected by automated decisions (Government of Brazil, 2023; Schertel Mendes & Kira, 2023).

For AI systems, and GenAI specifically, regulation must be approached democratically and on a global scale, with distinctive approaches to acknowledge that conditions vary between the Global North and Global Majority World. The key is to ensure that whatever institutions are in place at the global, regional, and national levels, they are used in ways that are consistent with human rights protection and the public interest (Souter, 2024). Allowing private technology firms and regulators primarily based in the Global North to set the parameters for AI regulation poses significant risks (Ross Arguedas & Simon, 2023), including the risk that choices about AI system design and deployment, made mainly in the Global North, will be inappropriate in different cultural, social, and political contexts.

There is also a risk that the largely hidden ability of Big Tech companies to exploit relatively cheap labour in the Global Majority World during AI system development will become more widespread if regulations do not take workforce extractivism into account (Monasterio Astobiza et al., 2022). Reframing debates about AI governance in Global Majority regions is crucial to acknowledge power asymmetries and to recognise that most "responsible" AI governance frameworks still allow powerful companies to "diffuse accountability, evade liability, and disregard rights" (Gurumurthy & Bharthur, 2023: 2), and it is argued that discussion typically excludes "marginalized communities and groups including women, racial and sexual minorities, small producers, workers, and Indigenous communities" (Gurumurthy & Bharthur, 2023: 3).

Scholars also point out that debates around requirements to ensure that AI systems are explainable are rooted in Eurocentric or Western perspectives. For example, in sub-Saharan African countries, where local informal savings and lending practices are common, AI tools to assess creditworthiness are found to exclude these practices (Effoduh, 2024). When the discussion turns to ethics, scholars in the Global Majority World point to bias towards the interests of the Global North, which then neglects approaches that differ from those adopted in Europe or the United States (Ali et al., 2024). The importance of differentiating between countries is emphasised by comparative studies of AI system governance, for instance, in Senegal and Cambodia (Heng et al., 2022). Generally, uniform approaches to governing AI systems are not likely to be viable (Ashok, 2023; Kakkar, 2023); yet, AI governance frameworks are being developed that transcend national boundaries and address the international implications of AI systems, which go beyond regional normative

approaches and global commitments to sustainable and accountable AI. So far, no organisation has succeeded in taking the lead in driving the development of AI systems in a way that is based on international solidarity and inclusive participation.

Privacy and data governance

The collection and processing of sensitive personal and non-personal data on an industrial scale by Big Tech companies requires the updating of governance to mitigate risks of privacy infringements and harms resulting from identity exposure. Specific rules apply for different types of data depending on their sensitivity and the risks associated with their misuse, with the aim of increasing transparency and accountability for data use (Kerber, 2020). In most countries, the collection and use of specific types of sensitive data – such as biometric data, health information, and personal identifiers – by digital service providers are subject to regulation (Payne et al., 2023).

The European Union's General Data Protection Regulation (GDPR) (EC, 2016b) refers to fundamental rights that are enshrined in its Fundamental Rights Charter: "Everyone has the right to respect for his or her private and family life, home and communications" and "Everyone has the right to the protection of personal data concerning him or her" (EC, 2000: Art. 7, 8). The GDPR is among the most comprehensive regulations. Implemented in May 2018, it sets out requirements for companies and organisations that collect, store, and manage personal data. It applies across all sectors but with specific provisions for some types of data-processing activities. For example, stricter conditions are imposed on the processing of special categories of personal data such as health information, biometric data, and data revealing racial or ethnic origin.

Compliance with the GDPR has proved challenging for the Big Tech companies. The demands for increased transparency and data handling accountability compelled Google to overhaul its privacy policies and practices. It revised its privacy policies to make them more understandable and accessible to users, simplifying the language and providing clearer explanations of what data is collected and how it is used. Its policies now include detailed descriptions of privacy controls that users can access to manage their personal information, aiming to ensure users have a better understanding and greater control over their data. More granular privacy controls were introduced in user account settings, and a Privacy Checkup tool was rolled out to guide users through privacy settings, aiming to help them make better informed decisions about their data (Houser & Voss, 2018; Murtaza & Salman, 2019; Waldman, 2020).

Despite these changes, whether the GDPR goes far enough to ensure data privacy is questioned (Lee, 2024; Wulf & Seizov, 2022). Tech companies are frequently charged with data breaches and unauthorised tracking of users. A

case brought under European Union competition rules against Alphabet for Google's alleged tracking of users by the European Center for Digital Rights (NYOB), a privacy advocacy group, resulted in scrutiny for unauthorised user tracking by its Chrome web browser (Halpern et al., 2024). Meta's privacy controls have been criticised for their use of so-called dark patterns: "tricks used in websites and apps that make you do things that you didn't mean to, like buying or signing up for something", creating opportunities for unauthorised data collection that are opaque to the user (Deceptive Patterns, n.d.; see also Brignull, 2023). Alleged privacy infringements can take a long time to resolve, as in the case of the Irish Data Protection Commission's investigation into whether Google uses sensitive personal data about race, health, or political preferences to target ads (Murgia, 2019; O'Faolain, 2024). The complexity of data collection practices and the volume of data pose significant challenges for achieving true transparency, and data-protection authorities struggle to keep pace with multiple cases before the courts.

The power asymmetry between Big Tech companies and their users can leave the latter unaware of the full extent and implications of data collection and data monetisation. The persistence of targeted advertising and misuse of data for political gain mean that the GDPR is not a panacea for human rights protection. Vigilance and a commitment to ethical data practice are essential to protect user privacy and maintain public trust. Nevertheless, the GDPR has served as a template (Bryant, 2021), influencing global data-protection standards with its comprehensive and extraterritorial reach, although approaches vary, reflecting unique sociopolitical and economic contexts. Research suggests that the GDPR has had a favourable impact on incorporating private rights, but this influence has been strongest inside the European Union. There is also evidence of a shift in tech company behaviour when data collectors are required to disclose data handling practices and undergo regular audits to ensure legal compliance (Linden et al., 2020).

In African countries, initiatives are being taken to introduce privacy and data-protection legislation, and there is ongoing review of data privacy laws (Abdulrauf & Dube, 2024). A data-protection law was put in place in Cape Verde in 2001, and by the end of 2023, 35 countries had enacted legislation and other measures were pending. However, in some cases, countries introduce exemptions for national security reasons and there are concerns about the robustness of the institutional protection of human rights (Andere & Kathure, 2024; Ndemo & Thegeya, 2023; South Africa Government, 2024). The African Union Convention on Cyber Security and Personal Data Protection (African Union, 2014) promotes a unified, continent-wide approach to cybersecurity and data privacy for 55 states and is seen as diverging from the GDPR in several ways (Andere & Kathure, 2024; Duncan, 2023).

Brazil's General Personal Data Protection Law (LGPD, *Lei Geral de Proteção de Dados*) mirrors the GDPR's comprehensive scope while incorporating elements tailored to the country's environment. The LGPD emphasises

principles of transparency, purpose limitation, and data minimisation, aiming to balance the protection of personal data with the facilitation of economic activities (Government of Brazil, 2018). China's approach to data protection, for example, the Cybersecurity Law and Personal Information Protection Law (PIPL), is tied to its broader strategy to balance economic growth and national security (He, 2023; US-China Commission, 2022; Voss & Pernot-Leplay, 2024), although Chin (2024) has argued that the main emphasis of the data laws is on data as a new factor of production and that this does not acknowledge epistemic rights.

In India, a Draft Personal Data Protection Act (DPDPA) was proposed after a landmark Supreme Court judgment – *Puttaswamy vs. Union of India* – in 2017, and passed in 2023. The DPDPA mirrors aspects of the GDPR and is aimed at “the processing of digital personal data in a manner that recognises both the right of individuals to protect their personal data and the need to process such personal data for lawful purposes” (Government of India, 2023a: 1). In some sectors, such as financial technology, sectoral regulations apply. The Reserve Bank of India plays an important role in regulating the financial services industry, establishing and enforcing self-regulatory guidelines, and penalising and suspending bank licences that do not comply with its data-protection guidelines and “know your customer” norms (Privacy Law Library, 2017; Punia et al., 2022). Under the GDPR, responsible parties are clearly defined as either data controllers or data processors: A data controller determines the purposes and means of processing personal data, while a data processor is responsible for processing data on behalf of the controller. This distinction is crucial for accountability, as it clarifies who is responsible for ensuring compliance and who will be liable if something goes wrong. In India, the DPDPA uses the concept of data fiduciary and data processor, where data fiduciaries are akin to data controllers under the GDPR and are tasked with exercising due diligence in the processing and securing of personal data.

Japan has updated its Act on the Protection of Personal Information (APPI) several times after its introduction in 2003, achieving GDPR compliance a year before the European Union's legislation came into force. Amendments have expanded the scope of individual rights, provided for stricter consent requirements, made data breach notifications mandatory, and limited the types of data that can be provided to third parties (Coos, 2022).

In the United States, the capacity of digital platforms to collect, process, and make data generated online available to third parties without user consent is subject to the federal Children's Online Privacy Protection Act (COPPA) of 1998 (updated in 2013), which governs the collection of information about minors, addressing issues of parental consent, confidentiality, and security, with safe harbour provisions and rules for data retention and deletion (US Congress, 2013). There is no single federal law governing data privacy generally, with recent bills (the American Data Privacy and Protection Act and the American Privacy Rights Act) failing to gain support. However, laws

do apply to data and telecommunications, health information, and credit, financial, and marketing information. And the Federal Trade Commission regulates to constrain unfair or “deceptive or unfair business practices and from unfair methods of competition” and takes action to enforce privacy laws (Kira et al., 2021).

The Health Insurance Portability and Accountability Act (HIPAA) includes data privacy and security provisions for safeguarding medical information and significant penalties for non-compliance (US Congress, 1996a). Companies specialising in healthcare data analytics aim to navigate the HIPAA by leveraging their technological expertise while capitalising on the commercial value of healthcare data. Palantir, for example, claims that its data management system and security protocols enable it to handle sensitive health information efficiently and securely. Thus, companies with substantial resources can thrive under data governance legislation by integrating compliance into their operational models. Smaller healthcare operators claim, however, that they are burdened by compliance costs (Chen & Benusa, 2017). Palantir’s expansion into the health sector and its uses of invasive surveillance software technology in the health sector and its security business have generated a robust literature on how companies are using data in ways that discriminate among different segments of the population (Lanzing, 2023; Ulbricht & Egbert, 2024). Methods for anonymising data are vulnerable to a range of de-anonymisation attacks, and it is unclear how methods compare when used for specific purposes (Sepas et al., 2022). The capacity to link data is built into the Internet, and the more sources of data that can be linked, the greater is the risk of de-anonymisation (Lane et al., 2014). In addition, there is the “tyranny of the minority”: where the volunteered information of the few can unlock the same information about the many because they happen to share the more easily observable characteristics that correlate with the revealed characteristic (Barocas & Nissenbaum, 2014a, b). While the HIPAA legislation is comprehensive in some respects, it does not address the complexities of new technologies, and the data generated by mobile health applications and wearable technology can fall outside the scope of this legislation.

Data protection to restrict access to abortion data came under scrutiny especially after the United States Supreme Court reversed the 1973 *Roe v. Wade* decision in its 2022 *Dobbs v. Jackson Women’s Health Organization* decision, which gave powers to the States to legislate on abortion. This has led to heightened concerns about the privacy and security of reproductive health data and the need to prevent data misuse and discrimination to ensure that individuals’ privacy rights are upheld (Dellinger & Pell, 2024; Roth, 2022). There are also state-level laws, including the California Consumer Privacy Act (CCPA), which is considered the most stringent. It requires businesses to disclose the categories and specific pieces of personal information they collect at or before the point of collection. It asserts the consumer’s right to know about the personal information that is collected and when it is

sold or shared with third parties. This law broadened the concept of data sale, potentially encompassing many types of data transactions not typically considered sales and requiring businesses to reevaluate their data practices (US State of California, 2018).

Approaches to data governance differ globally, with some being tailored to address inequalities. This makes it important to differentiate between approaches in the higher-income and middle- and lower-income countries (Sampath, 2021). Data protection and privacy legislation may share common themes, such as provisions for consent and attention to data subject rights, but diverge in areas concerning requirements for data localisation and other features. For instance, the GDPR mandates strict consent requirements and offers robust data subject rights. In contrast, the CCPA, while also emphasising consumer rights and transparency, introduces a unique private right of action and provisions tailored to California's legislative context. India's DPDPA incorporates stringent data localisation requirements emphasising digital sovereignty, local data control, and geopolitical considerations.

All these legislative approaches shape data privacy and individual rights protection and the standards that are adhered to (Park, 2020). Governing data is a delicate endeavour, involving a struggle to balance the need for data security and privacy with the benefits of data for sector applications (e.g., health, finance, environment monitoring, or monetisation). By implementing specialised governance measures, enhancing transparency, and promoting public awareness, the aim of governance measures is to safeguard sensitive personal data against misuse. However, with diverse parties involved in data collection – from tech giants and startups to governments and third-party contractors – there are unique challenges in achieving governance that assures accountability, fairness, and transparency in how data is collected and used (Dolata et al., 2022).

The legislative tools such as the GDPR, the AI Act, or the CCPA provide foundational frameworks for comprehensive protection for individual privacy. They set benchmarks for global regulation intended to empower consumers so that they gain more control over the personal information that businesses collect about them – including transparency about data use and the right to delete collected personal data. The aim of these interventions is to ensure that data-driven products, especially those employing AI systems, involve the ethical collection and use of data to ensure transparency and mitigate algorithmic bias in a way that prevents discriminatory and harmful outcomes (Mitchell et al., 2021). However, evidence indicates that people's rights are inadequately protected. One reason is the global nature of cross-border data transfer operations, which confront the jurisdictional complexities when data handled overseas must be accorded the same level of protection as within an originating country (Reis et al., 2024). This is especially difficult when the operations of digital platforms weaken or

compromise a country's privacy and data-protection practices to enhance user engagement and advertising revenue (Floridi, 2013).

The mere existence of data-protection laws does not translate automatically into effective enforcement. The availability of adequate resources, financial and human, varies by country, even though these resources are crucial if data-protection authorities are to perform their functions (Bennett & Raab, 2006). There are major concerns about how robust existing privacy and data-protection legislation is in practice, with numerous examples of human rights abuses when data are used, for example, to support systemically biased facial recognition systems to identify individuals participating in crime, including riots, using predictive policing, or, for example, to identify migrants' children (Chouliaraki & Georgiou, 2022; Lanzing, 2023).

Governing mis- and disinformation

All the foregoing governance measures have an impact on incentives to produce and circulate mis- and disinformation and on efforts to curtail the production and spread of this information. It is crucial to distinguish between governance that supports human rights. Government censorship using technical measures that violate privacy can also result in blanket indiscriminate censorship (Berger et al., 2023). Too often,

regulation imagines itself as simply enforcing the given and natural norms of a decent society [...]; but from a sociological or anthropological point of view we know that law is always actually enforcing the mores of the dominant group that controls the content of law. (Post, 2009: 130)

AI systems remain, to a certain degree, unregulated, despite the introduction of legislation, and this presents challenges and opportunities for the sustainable design of information ecosystems that support information integrity. The United Nations AI Advisory Board has stated that voluntary commitments to mitigate the harmful effects of AI generated content “that is likely to deceive in a way that causes harm or societal divisions, or which promotes war propaganda, conflict and hate speech [...] *do not sufficiently mitigate societal risks* [emphasis added]” (UN, 2024a: 35, 36), and the board has called for a global, multi-stakeholder approach, alongside binding commitments. Since these systems are always used in social contexts, additional regulation within these contexts usually applies, as discussed above.

Freedom House research identified 22 countries with legislation mandating or providing incentives for Internet platforms to use AI tools to eliminate online speech that the state deems undesirable (Funk et al., 2023). Thus, both authoritarian and some democratic countries are using AI systems to broaden and reinforce censorship. For example, chatbots in China are programmed not to react to inquiries about Tiananmen Square. YouTube and X were required by the Indian government to restrict access to a documentary that

showed the violence that occurred when Prime Minister Modi was Gujarat's chief minister (Ryan-Mosley, 2023). International standards, regional norms, and national laws need to be intertwined to effectively meet the challenges of automated content governance systems (von Ingersleben-Seip, 2023). The soft-law character of many existing rules (such as commitments to limit mis- and disinformation) makes it difficult to assess the effectiveness of AI governance. In addition, the absence of robust AI content classification systems enables state and commercial actors to exploit the capacity of personalisation systems to prioritise engagement-rich content (Thomas, 2022).

Campaigns against mis- and disinformation by civil society organisations can support causes that are pro-government or attack opposition political leaders. When the ambition is to distract, suppress, or polarise opinion (Bradshaw & Howard, 2019; Richter, 2024), the affordances of information ecosystems can be pro- or anti-democratic, and governance may favour legislation that does not uphold or defend human rights (Bhatia et al., 2023; Borelli, 2023). Online activity can support protests against authoritarian regimes, just as it can lead to surveillance when trace data is generated that bolsters the interests of authoritarian states (Mutsvairo & Rønning, 2020). The framing of the issues and measures to mitigate harms linked to mis- and disinformation depend on “political decisions about normative issues” (Erman & Furendal, 2022: 267), reflecting the interconnected nature of digital technologies and societal norms. Governance rules are means of exercising a normative influence over information ecosystems. Initiatives to mitigate the harms of mis- and disinformation may be framed by international human rights, but the best means of institutionalising rights protections are contested (Roberts et al., 2024).

Between 2011 and 2022, 78 countries passed sector-specific laws designed to limit the spread of online mis- and disinformation. Legislation may focus on improving the transparency of the datafication process and the accountability of Big Tech companies, increasing media and information literacy (MIL) and AI literacy, or both. In some instances, the emphasis is on criminalising the creation and distribution of content, which, in authoritarian states, paves the way for state evaluations of what constitutes “fake news”, leading to abuse and undermining of freedom of expression, including press freedom (Lim & Bradshaw, 2023). Evidence on whether the implementation of legislation in this area adheres to human rights principles is mixed (Rajkumar & Ashraf, 2023). A large-scale review of research published from 2020 to 2022 found that many countries have laws to govern online content, but these are often vague and lead to arbitrary actions by public authorities and platform operators. There are also gaps between expectations and how local issues are dealt with when rights protection and business model goals are in tension (Berger et al., 2023). The large platform companies support the United Nations Guiding Principles on Business and Human Rights (Ruggie, 2011) noted earlier, but their practices often fall short. In addition, multilevel

governance can mean uneven rights protections and struggles to coordinate multiple initiatives (Saurwein & Spencer-Smith, 2020; Hatfield et al., 2023).

The definitions used in legislation on mis- and disinformation and penalties for distributing it vary substantially, based on a study of 32 countries (mostly in the Global Majority World with 12 designated as democracies, 19 as authoritarian, and 1 unclassified). There is an absence of clearly designated authorities to determine what constitutes information to be targeted, and much legislation was evaluated as risking “*more harm than good* [emphasis added]”, partly because nearly half of policies gave the government the authority to decide what is “fake news” (CNTI, 2024: 1). Legislation in this area also draws on laws dealing with defamation, sedition, technology regulation, cybercrime, and online harassment that have been in place since colonial times, with potentially detrimental implications for protecting the right to freedom of expression.

The European Union introduced dedicated measures to counter mis- and disinformation with an Action Plan (EC, 2018) centred on improving capabilities to detect, analyse, and expose mis- and disinformation, with the aim of strengthening coordinated responses, mobilising the private sector, raising awareness, and increasing societal resilience. A voluntary Code of Practice on Online Disinformation was updated in 2022 (EC, 2022e) and then formally recognised in the Digital Services Act in 2025. The code commits industry to address mis- and disinformation, political advertising, and the integrity of services, and aims to empower consumers and the research community (EC, 2018). A monitoring mechanism – the European Digital Media Observatory – was established as a hub for fact-checkers and for those studying mis- and disinformation. Early signatories of the code included Meta, Google, then Twitter and Mozilla, parts of the advertising industry, and then Microsoft and TikTok (Saurwein & Spencer-Smith, 2020).

The European Union’s AI Act places an obligation on providers of the largest online platforms or online search engines “to identify and mitigate systemic risks that may arise from the dissemination of content that has been artificially generated or manipulated, in particular risk of the actual or foreseeable negative effects on democratic processes, civic discourse and electoral processes”, including through disinformation (EC, 2024c: para 120). The act specifically mentions systemic risks through “the facilitation of disinformation or harming privacy with threats to democratic values and human rights” (EC, 2024c: para. 110). This aligns with the DSA (EC, 2022c), although under the DSA, platforms (intermediary services) are not liable for the content they host if they act as a “mere conduit”, provide temporary storage caching, or host illegal content without knowledge and act quickly to remove or disable it when they obtain such knowledge. The European Union’s Code of Practice on Online Disinformation is not without critics regarding what is defined as mis- or disinformation, whether the measures are disproportionate, or whether they amount to tinkering with problems while

data monetisation models are allowed to stand (Culloty, 2021; Monti, 2020; Nenadić et al., 2023; Pamment, 2020). Nevertheless, European legislation has inspired other countries to take similar action, with liability provisions for the information the digital platforms host varying around the world (DiGi, 2022; Wilding, 2021).

With much broader country scope, the Council of Europe's (2022: para. 4.1) recommendation on media coverage of election campaigns states that,

the algorithms used by public and private actors to rank and display political advertising and electoral communication material, and those used in content moderation practices, should be transparent and verifiable, especially regarding potential bias and inaccuracies of the systems used.

Platforms are recommended to “act against misrepresentation and the intentional spread of political disinformation, while ensuring full respect for the rule of law and human rights standards [...] notably the right to freedom of expression” (Council of Europe, 2022: para. 4.2).

In the United States, diverse interpretations of First Amendment speech rights and protections have led to controversy around the need for content governance (Forum on Information and Democracy, 2024d). The Big Tech owners of digital platforms benefit from Section 230 of the Telecommunications Act of 1996 (US Congress, 1996b), which exempts platforms from liability for content they host on their platforms. Providers or users of interactive computer services are not treated as publishers or speakers of any information provided by another content provider. They therefore have broad immunity from liability for the content they host, and there is ongoing controversy about whether this should change (Cramer, 2020). Legislative proposals aimed at curtailing the circulation of content deemed to be harmful typically fail to attract sufficient congressional support and debate about how platform immunity might be circumscribed is highly politicised, being met with free market arguments and the claims that competition will eliminate problems (Flew, 2021).

In sub-Saharan Africa, a study of 46 countries indicated that 20 pieces of legislation against disinformation had passed under penal codes, two countries had legislated against falsehoods, nine countries had addressed disinformation, focusing on the intent to mislead or deliberately spread false information, and eight were specifically penalising misinformation (Sey et al., 2022: 159). In Asia, critics of measures to combat mis- or disinformation argue that these are used as “a cover to silence dissent and legitimate criticism of the state” (Jayasinghe et al., 2022: 192). When digital platforms provide spaces for political activism – including by opposition parties as in Southeast Asian states, state authority is maintained through political pressure and Internet controls (Sinpeng, 2020).

In Brazil, the government's policy of leaving it to platform companies to moderate election-related content was found to have led to the use of opaque

criteria to detect inappropriate content, with platform decisions having an advertising effect and endorsing specific candidates (Santini et al., 2023). Delayed initially by lobbying by Big Tech companies, the Brazilian law on AI calls for the development, implementation and use of AI to support the integrity of information by protecting and promoting the reliability, accuracy and consistency of information (Government of Brazil, 2024).

In India, the 2021 Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules were updated in 2023 to regulate digital media through the Ministry of Electronics and Information Technology (MEITY). This applies to social media intermediaries, news publishers, and the Big Tech owners of digital platforms. Faced with a severe backlash from civil society and digital rights organisations, parts of the legislation relating to fact-checking were suspended by the Supreme Court (Jain, 2023; Government of India, 2023b; Oladapo & Ojebode, 2021; Roy, 2024). In Nigeria, government strategies were seen as threatening to shut down the media space to silence the voices of critics of the government using the 2019 Protection from Internet Falsehoods and Manipulation and Other Related Matters Bill, which criminalised sharing information that diminishes public confidence in the government (Oladapo & Ojebode, 2021).

Other countries' legislation is criticised for failing to uphold the right to freedom of expression and for intrusive surveillance: for example, in Malaysia, the Anti Fake-News Act (Malaysia Government, 2018); in Singapore, the Protection from Online Falsehoods and Manipulation Act (Singapore Statutes, 2019); in Russia, “fake news” laws (Richter, 2019); in Bangladesh, the Cyber Security Act (Government of Bangladesh, 2023); and in China (Reuters, 2019; Repnikova, 2018).

While new methods of generating and circulating mis- and disinformation using AI tools present a threat to political processes, strategies to combat this risk infringing human rights due to the use of vague definitions and, in some cases, excessive and disproportionate sanctions (UN, 2023b). The effect of some laws is to position either digital platform owners or states as the arbiters of “truth”, which can lead to abuses of basic freedoms, and co-regulatory approaches are seen as being best positioned to reduce the risk of such abuses (Dittrich, 2019). It is also argued that policymakers tend to focus myopically on digital platforms and social media when they tackle mis- and disinformation. This ignores the roles of the Internet and other intermediaries, including Domain Name System registries, financial intermediaries, and Internet certification authorities in spreading this information. Greater attention should be given to these actors if mis- and disinformation are to be tackled effectively (Bradshaw & DeNardis, 2022). Others argue for more attention to restoring trust in journalism in response to mis- and disinformation instead of government-mandated Internet shutdowns or other measures that suppress online expression, especially during elections (Das & Schroeder, 2021; De Gregorio & Stremlau, 2021).

The evidence confirms that it is crucial to attend to how governance approaches differ. Much of the literature on information ecosystems governance is focused on the United States, Europe, and, increasingly, China, although alliances among the BRICS countries are working to establish digital sovereignty (Jiang & Belli, 2025). Clear lines of accountability are needed to ensure that those responsible for digital risks and harms can be held accountable and that people's rights to freedom of expression, privacy, and autonomy are protected.

Safely navigating information ecosystems

People have varying abilities to identify mis- and disinformation and to respond to it in ways that protect adults and children from harm and enable them to participate in informed public debate based on accurate information. Media and information literacy (MIL) and AI literacy initiatives are either mandated through government legislation or encouraged to improve the public's abilities to critically assess and interpret online information and, specifically, to detect mis- and disinformation and exercise agency in their choices about how they respond.

Media and information literacy

MIL initiatives aim to empower people to manage their online activity, patterns of information consumption, their capacity to identify and to protect themselves from harms linked to mis- and disinformation, and other issues (Friesem et al., 2022). MIL training aims to,

help realise human rights and facilitate human flourishing – including diverse forms of creativity, human connection, community and political participation – the institutions and structures of our society must make room for people's agency, knowledge and self-determination, finding ways to recognise and value and enable these, perhaps transforming themselves in the process. In short, media literacy is not a stand-alone project. To see the positives of media literacy, we have to imagine a positive vision of society – what it could be, what people want it to be, what they need it to be. (Livingstone, 2023: para. 23)

MIL initiatives aim to provide competences and skills to navigate through information ecosystems in which information integrity and the material conditions of people's lives vary enormously. National, supranational, and civic initiatives seek to enhance skills and position people to make informed choices. States may have obligations to encourage MIL, as in the European Union and the United Kingdom (Durach et al., 2024; EC, 2017; Frau-Meigs & Corbu, 2024; Pentney, 2024), but attention to this area and resources vary within and between the Global North and the Global Majority World.

MIL is addressed by extensive research (see De Abreu, 2022; Frau-Meigs, 2024b; Frau-Meigs et al., 2017; Haider & Sundin, 2022; Livingstone & Blum-Ross, 2020; Ofcom, 2024b; Ragnedda & Gladkova, 2020; Schmarzo, 2023; UNESCO: n.d.).

Media literacy may refer to “the ability to use, understand and create media and communications in a variety of contexts” (Ofcom, 2023b: 3; see also Jeong et al., 2012; Livingstone, 2008). Media literacy is also defined as “the ability to access, analyze, and produce information”, the fundamental objective being “critical autonomy in relationship to all media” (Aufderheide, 1993: 1). Information literacy refers to the ability to find, evaluate, and proficiently use information, and some prefer to focus on digital literacy in the context of equity and inclusion (De Paor & Heravi, 2020; Menon, 2017, 2024).

MIL conceptual frameworks acknowledge the essential role of literacy in influencing how people construct their identities and realities through online interactions and relationships (Cho et al., 2024). Definitions of MIL are changing over time to reflect new patterns of media and information presentation and consumption (Wuyckens et al., 2022), and there is a variety of methodologies for assessing the impact of MIL initiatives (Frau-Meigs, 2022). Interventions aim to equip people with key competences, for example, to understand the importance of information and the ability to think critically when engaging with information and to find reliable information (Grizzle et al., 2021; Jones-Jang et al., 2021; Vuorikari et al., 2022), and there are numerous toolboxes for combating mis- and disinformation (Kozyreva et al., 2024; da Silva et al., 2019; Dadakhonov, 2024) (see Figure 5.3).

Figure 5.3 Non-exhaustive MIL competencies



SOURCE: Frau-Meigs, 2024c: 4; redesigned by Mattieu Mercier and adapted by Nordicom.

Attention may focus on helping online users to acquire fact-checking competences. This may involve lateral reading skills or taking independent steps to verify information. MIL initiatives may be implemented through school curricula or online (e.g., online videos, pop-ups, online games). Some are informed by the theory that “inoculating” people by sensitising them to inaccurate information can help them to recognise common mis- and disinformation formats or tactics, but it is acknowledged that this only works in certain circumstances and for certain people, and there is no universally effective solution (Kozyreva et al., 2024). Among these kinds of interventions are accuracy prompts, prebunking and debunking, creating friction, encouraging lateral reading, providing media literacy tips, offering rebuttals to science denialism, self-reflection tools, and learning about social norms, as well as providing warning and fact-checking labels (Kozyreva et al., 2024). In the case of lateral reading or thinking, this is often associated with education that aims to develop critical consciousness through education (Freire, 1974; Wineburg & McGrew, 2019). Adjin-Tettey (2022) found that experimental one-off MIL training made respondents more likely to be able to detect the authenticity of information. These interventions tend to be more audience-centred, with less attention to the use of technical skills.

Studies of the effectiveness of MIL interventions overrepresent experience in the Global North, but they do indicate that better Internet skills are associated with thinking more about the credibility of online information in the case of young people (Livingstone et al., 2023a). Although evidence on the effectiveness of MIL strategies is absent for many countries, both teachers and learners need to be included in acquiring values and practices consistent with respect for others, and “whole-school” approaches are needed to ensure that curricula revisions are culturally appropriate.

Research on the relationship between children’s digital skills and outcomes has shown that better access to digital technologies at home or at school is linked to more positive attitudes to digital technology. Although there is little fine-grained research on children’s awareness of algorithms, the roles of their use of digital platforms in their lives, and how their rights are affected, digital skill levels are found to be positively linked to children’s online behaviours (e.g., privacy protection, deleting messages, or blocking and dealing with cyberbullying and sexual images). Skill levels are associated with whether they are likely to participate in boycotts, rallies, or online campaigns, but in the case of young adults, skill levels are not found to influence voting and other forms of democratic participation (Livingstone et al., 2023a; see also Livingstone et al., 2024).

In addition, civic education is found to play a role in MIL by increasing political efficacy and is a predictor of interest and trust in news. Students participating in a civic education programme in the United States, in a study in the early 2000s, reported greater self-efficacy, an effect carrying over to increased political attentiveness and knowledge of candidate positions, with political attentiveness increasing knowledge and voting (Pasek et al., 2008). In emerging democracies, civic education has also been found to have favourable effects on levels of political information and participation, including a reduction in authoritarian nostalgia and an increase in desired political behaviour (Finkel, 2014; Finkel et al., 2024). In experimental research with participants who were supportive of democracy before the experiment in post-Soviet countries (Belarus, Moldova, Russia, and Ukraine), it was found that civic education programmes led to an increase in young people’s support of democratic institutions, democratic attitudes, and perceived political efficacy, albeit with small effects (Pospieszna et al., 2023). Qualitative research in Europe on how young people respond to targeted political advertising has suggested that literacy programmes may help them to detect micro-targeting online tactics (Hirsch et al., 2025).

Positive outcomes of media literacy interventions have been found in relation to knowledge, criticism, influence, realism, beliefs, attitudes, norms, self-efficacy, and behaviours, concluding that interventions were generally effective. However, the caveats were that effects might be greater for outcomes

related to knowledge and realism than for attitude and behaviour outcomes. Interventions were more likely to be successful if reinforced over multiple sessions (Jeong et al., 2012).

In Cote D'Ivoire, Facebook's most widely shared posts were about assassination rumours, vaccine scepticism, xenophobic hate speech, and doxing of political opponents. When literacy training was introduced, the intervention did not change how people consumed or shared information, but the importance to online users of affirming group identities was found to take precedence over whether information was misleading. In Mali, Burkina Faso, and Niger, efforts to tackle online gender violence against women and girls illustrated the importance of collaborations between organisations that specialise in information verification and those involved in humanitarian and development initiatives to combat the flow of sexist information. And in Cameroon, a strategy to promote media and information literacy through awareness-raising for women emphasised the need to recognise power structures if inclusivity was to be achieved, and repressing harmful content was difficult to achieve (Birwe, 2024; NDI, 2023; Zibi Fama, 2024). Research on gender-based violence in the Central Sahel region has emphasised the need for education to tackle the circulation of harmful content on social media (Uyheng & Carley, 2024).

In Sri Lanka in 2021, when 57 per cent of those aged 5–69 were deemed digitally literate, the definition of literacy was revised to align with UNESCO's global standard to capture computer literacy and literacy in the use of smartphones, the primary way that people at the “bottom of the pyramid” access the Internet (Fonseka, 2024; UNESCO, 2022a). In China, an Education Informatization 2.0 Action Plan emphasising wisdom education is said to have shifted investment from training in information technology skills towards improving student and teacher information literacy, with a focus on all-round human development (Yan & Yang, 2021).

Literacy training cannot compensate for a failure to ensure that the design of technologies is safe before services are deployed in the market. It is also essential to recognise that literacy training can result in improved skills to produce and circulate hateful, dehumanising, and violent content (Banaji & Bhat, 2022). For this reason, education and literacy initiatives should come before experimentation and deployment of new technologies. As an expert participating in our research assessment observed, having largely failed to curtail the spread of online mis- and disinformation using legislation and MIL initiatives, attention is turning to AI literacy training.

AI literacy

Encounters with mis- and disinformation are shaped by the AI systems that generate content and personalise use algorithms to personalise, moderate, and curate information flows. Research focuses on the contribution that AI literacy can make (Okunlaya et al., 2022), with some scholars preferring the term algorithmic literacy, a term in use since 1985, when it was defined as the ability of people without technical training to recognise when they are interacting with a system driven by algorithms, to reason about what kinds of data might be collected, and to respond based on their decisions about how they want to interact with these systems (Boots et al., 2024; Oeldorf-Hirsch & Neubaum, 2025). One of the most cited definitions of AI literacy is “a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace” (Long & Magerko, 2020: 2). Whereas MIL typically refers to competencies required to use information and to communicate (Henderson & Corry, 2020), AI literacy is likely to include the ability to comprehend the core principles and concepts of AI systems. Much effort has gone into defining what these competencies require in terms of specific skills. Training is expected to enable people to answer questions (Long & Magerko, 2020), such as:

- What is AI? (e.g., knowing differences between AI and other digital technologies.)
- What can AI do? (e.g., knowing what these differences mean for how AI can be used and its strengths and weaknesses.)
- How does AI work? (e.g., understanding the principal technical elements of AI.)
- How should AI be used? (e.g., understanding the ethical issues raised by use of AI.)
- How do people perceive AI? (e.g., understanding common misconceptions about AI; making sense of AI; trustworthiness of AI.)

These competencies exemplify a broad view of AI literacy. AI literacy research suggests a more specific set of skills or micro competencies, as compared with MIL, which encompasses media, document, and data competencies.

**Media competencies:**

- Knowing the new context of news production and amplification via algorithms.
- Being suspicious and aware of weak signals of disinformation.
- Fighting confirmation biases and other cognitive biases.

Document competencies:

- Setting limits to tracking to reduce targeting (as fewer data are collected from devices).
- Browsing anonymously (e.g., use of virtual private networks, or VPNs).

Data competencies:

- Paying attention to platform adherence to data-protection rules.
- Mobilising for more transparency and accountability about the impact of data use.
- Signalling or reporting to platforms or web managers if data misuses are detected.
- Commenting and/or rectifying “fake news”, whenever possible.
- Alerting fact-checkers, journalists, or the community of affinity.

(Frau-Meigs, 2024a)



These competencies are expected to encourage a more proactive and potentially empowering approach to combating mis- and disinformation (Wang et al., 2023). In other words, they are not just about coping with mis- and disinformation, but rather boosting competencies for checking for possible exposure to mis- and disinformation (e.g., be suspicious and aware of weak signals) and competencies designed to reduce the risk of exposure occurring in the first place (e.g., set limits to tracking so as to reduce targeting, including ensuring fewer data are collected from devices, using virtual private networks, or contributing to data governance).

AI literacy as a tool for empowerment and resistance to the control that AI systems exercise over people’s activities is commanding some attention (Stamboliev, 2023). Practical solutions are essential to address people’s lack of awareness about how algorithms make decisions that inhibit their agency. Bottom-up solutions are needed to respond to the “challenge of algorithm opacity by looking at the end-user (not the producer) and empowering citizens to analyse algorithms critically and creatively, in the hope of bringing insights in their own information consumption” (Frau-Meigs, 2024a: 512).

Empowerment and resistance depend as much – if not more – on paying attention to the inputs of AI systems as to their outputs. Hence, the term “data literacy” is used to refer to the key set of competencies that people need to exercise control over their personal data, including what they allow to be collected and with whom it is shared. This is reflected in UNESCO’s definition of AI literacy as a combination of algorithmic literacy and data literacy:

[AI literacy] comprises both data literacy, or the ability to understand how AI collects, cleans, manipulates, and analyses data; and algorithm literacy, or the ability to understand how AI algorithms find patterns and connections in the data, which might be used for human-machine interactions. (UNESCO, 2022b: 11)

Government bodies as well as international and civil society organisations are making concerted efforts to promote AI literacy. For example, the governments of China (China Ministry of Education, 2019), Germany (German Federal Ministry of Education and Research, 2021), India (Stanly, 2024), the United Arab Emirates (UAE, 2018), and the United States (US Government, 2020) have put initiatives in place, and there are measures throughout South-East Asia and the European Union (see Laupichler et al., 2022). The focus of these programmes tends to be on integration within primary, secondary, and tertiary education. It is vital, however, that the needs of older people are provided for, with programmes taking account of evidence that they are generally less digitally literate and thus start from a lower base (Loos & Ivan, 2023; Moore & Hancock, 2022). AI literacy programmes must keep pace with rapid advances in AI technologies now that major search platforms are using GenAI to provide summarised responses to search queries. The appeal for users is no longer having to examine the list of links that a traditional Internet search produces, but the challenge becomes assessing the quality and relevance of the results with respect to the answers they are looking for, with the risk that these summaries provide an inaccurate or even false (hallucinated) representation of the sources summarised (Maynez et al., 2020).

GenAI tools such as ChatGPT and DALL-E are the result of an incremental process of technological innovation, and it may be argued that regulation to combat mis- and disinformation should not be introduced if there is a risk that this will slow the pace of technological change (Ross Arguedas & Simon, 2023). AI literacy training offers an attractive option to those who resist regulatory intervention, especially if it is based on shared competency development and assessment criteria and is effective in producing innovations in GenAI tools that operate in ways that are consistent with democratic processes (Ng et al., 2021). Evidence of the impact of AI literacy training on interactions with and consumption of online information is limited. There is evidence, however, of growing awareness of potentially detrimental impacts of uncontrolled data collection on individual privacy that may be linked to a rise in AI and data literacy.

An evidence assessment of research on media literacy conducted between 2011 and 2021 with a focus on misinformation found that “strategies that prompt conscious and rational engagement with content and develop critical thinking skills” were effective, though variably so, and that future research was needed using more robust and comparable methodologies and including vulnerable and marginalised populations (Anstead et al., 2025: 2129). In addition to the practical challenges of generating and circulating information, producing less biased datasets, and understanding how algorithms work, individuals need to be able to make sense of information. When the critical literacies of populations are developed, this is likely to influence decisions about whether there is “equality and/or symmetry between human and non-human actors, and [... the] conceptualization, development and understanding of new forms of intelligence we would like to live with in the future” (Jandrić, 2019: 35; see also Ritzer et al., 2024). Critical literacy skills are likely to encourage imaginative approaches to the design and use of AI systems, addressing ethical issues and helping to recognise when AI tools, data collection, and their use are helpful and when they are not (Deuze & Beckett, 2022).

Conclusion

This chapter has emphasised the duty of states to not violate human rights and fundamental freedoms as well as to protect them. We reviewed legislative developments including direct state regulation and co-regulatory initiatives as well as industry self-regulation to highlight variations in the normative goals and rules put in place to govern the components of information ecosystems. The analysis gave particular attention to the diverse approaches to balancing the interests of governments, Big Tech companies, and individuals and groups within society. We then examined the principles and voluntary commitments that are being agreed on at a global level and then how these are translated into policies for governing the network infrastructure layer of information ecosystems, digital platforms, news media, AI systems, and data privacy.

Particular attention was given to the legislative framework that has been put in place by the European Union, highlighting the features of the risk mitigation frameworks of the DSA, AI Act, and Media Freedom Act, as well as other pieces of legislation, comparing these with the numerous legislative interventions in other countries and regions. We particularly highlighted the risks associated with news media regulation when it negatively impacts press freedom and the need to tackle structural inequalities within the news media industry, including the dependencies of news media companies on digital platforms. In the case of AI system governance, we emphasised disparities in decision-making between the Global North and the Global Majority

World and the need for distinctive governance approaches to enhance the transparency and accountability of the Big Tech developers of AI systems. We also highlighted evidence indicating that data governance arrangements inadequately protect rights to privacy.

This chapter examined research on governance initiatives specifically intended to tackle mis- and disinformation, drawing attention to the use of diverse definitions and penalties that position states or Big Tech companies as “arbitrators of truth” with the authority to suppress information in untransparent ways and in ways that compromise people’s rights and freedoms or allow intrusive surveillance. In the final section, we examined MIL and AI or data literacy initiatives, emphasising the need to encourage critical engagement with online information and the fact that AI literacy training is needed at all stages of AI system development and deployment. We also emphasised that literacy training on its own cannot compensate for technology embedded in information ecosystems that is not designed to keep people safe or to facilitate their informed engagement in the public sphere.

Governing information ecosystems in practice

Introduction

Mis- and disinformation present major challenges for governing institutions. Legislative frameworks and institutional arrangements are being put in place or strengthened in countries and regions around the world to govern components of information ecosystems, as we discussed in Chapter 5. In this chapter, we look at whether these governance frameworks encourage practices that are effective in tackling harms linked to mis- and disinformation and strengthening the integrity of information. The chapter begins with research aiming to understand how well the public understands the risks and harms associated with mis- and disinformation and their views about various approaches to content moderation. We also examine research on policymakers' understanding of AI systems and digital platforms. We then turn to measures that are taken to combat mis- and disinformation using content moderation strategies, looking first at measures to provide the public with information and the activities of fact-checking organisations. Next, we look at government actions, including the use of codes of practice, and at Big Tech self-regulatory measures.

In the last main section, we examine alternative data governance practices – individual and collective – that aim to achieve data justice by resisting the exploitative business models of Big Tech companies. The need for meaningful democratic discussion and action to address power imbalances within information ecosystems is emphasised. A concern throughout this chapter is whether rights to freedom of expression, privacy, and equitable treatment are jeopardised by efforts to curtail online mis- and disinformation – that is, whether the stated aims of governance rules are falling short of normative expectations when they are either ignored or implemented in ways that do not align with democratic values and rules.

Awareness of mis- and disinformation

Governments have a positive obligation to protect human rights, as indicated in Chapter 5. This creates a responsibility to ensure that people are aware

of the risks they face while navigating information ecosystems. We discussed media and information literacy (MIL) and AI literacy initiatives that can help to fulfil this obligation in Chapter 5. In this section, we examine what is known about both adults' and children's awareness of risks and harms and their abilities to discern the difference between mis- and disinformation and accurate information.

Public awareness of risks and harms

Research on public awareness of the problems created by mis- and disinformation can be used to inform those with a mandate to govern information ecosystems about priorities for action. Levels of “unfreedom” in relation to the news media in many countries are high, with 85 per cent of the world's population seeing a decline in press freedom over the five years leading up to 2022 (UNESCO, 2022d). In view of restraints on what is and can be expressed, we might expect publics to be developing a heightened awareness of how online mis- and disinformation contribute to these restraints and to have views about what steps governments or Big Tech companies should take to counter problems associated with declining information integrity (Advox Team, 2024). The evidence shows a varied picture.

Research on awareness of risks and harms linked to mis- and disinformation varies substantially across countries. A study including 142 countries with over 150,000 participants found that, globally, the majority of regular Internet users (59%) reported being worried about misinformation, with young and low-income groups being most likely to be concerned. Perceptions of risk were highest in the Latin America and the Caribbean regions (74%) and lowest in South Asia (31%). Differences were not related to the prevalence of misinformation, but concern was highest in countries with liberal democratic governments (Knuuttila et al., 2022).

In another global survey with 8,000 participants aged 19 years and over in countries that were to hold elections in 2024 (Algeria, Austria, Bangladesh, Belgium, Croatia, Dominican Republic, El Salvador, Ghana, India, Indonesia, Mexico, Romania, Senegal, South Africa, Ukraine, and the United States), 89 per cent of interviewees agreed that “governments and regulators should be able to require social media platforms to put in place trust and safety measures during election campaigns to protect the integrity of elections”; 85 per cent reported concern about the impact of disinformation; and 78 per cent indicated that they had read content online that had been deliberately falsified (Ipsos & UNESCO, 2023: 8).

The Reuters Institute's global large-sample survey on digital news, encompassing Africa, Latin America, the Asia-Pacific, and Europe, found that despite relatively strong awareness about how information is managed on digital platforms, concerns about discerning what is “real” and what is “fake” news vary by region (Newman et al., 2023: 10). In the United Kingdom, 60

per cent of social media users were confident in their ability to identify a fake social media profile. Although 77 per cent reported thinking about whether online information was truthful, there was a high risk of mistaken judgments and of being misled (Ofcom, 2023a). In Europe, older people are generally found to be more likely to share mis- or disinformation, echoing similar findings in Canada and the United States (Frau-Meigs, 2022; Hunsaker & Hargittai, 2018; Jung & Sundar, 2016; Moore & Hancock, 2022; Schreurs et al., 2017). Abilities to detect whether texts are authored by a human also vary considerably (Dugan et al., 2023).

People use a variety of competences to discriminate between trustworthy information and mis- and disinformation. Interviews with participants selected for their reported lack of trust in their countries' news organisations in Brazil, India, the United Kingdom, and the United States addressed how people make sense of information on digital platforms (Facebook, WhatsApp, and Google) and the methods they use to detect mis- and disinformation. These included mental shortcuts such as the presence of visuals, brands, headlines, and advertising sponsors. Social cues (e.g., who shared the information) and platform cues (e.g., likes or labels) were used to assess trustworthiness, which varied with the affordances of the platforms (Ross Arguedas et al., 2022c). Research also has examined whether people can distinguish between LLM- or human-produced tweets, finding that they cannot and that as their confidence in their ability to do so decreases, they become resigned and may simply give up on making an effort (Spitale et al., 2023).

A study with Spanish participants explored the “nobody-fools-me perception”, that is, people’s overconfidence in their individual abilities to detect mis- and disinformation (in this case, concerning the Covid-19 pandemic) coupled with a self-belief that they are more immune to such information than others. Informed by concepts of self-perception, self-efficacy, confirmation bias, miscalibration, misplacement, and misestimation, younger people were found to generally believe that older people were more likely to be fooled by mis- and disinformation and that younger people were less likely to fact-check. Those with higher levels of education were more confident about their ability to detect mis- and disinformation (Martínez-Costa et al., 2023). Research also shows that perceptions of immunity to deception tend to increase during periods of high information exposure (Tang et al., 2021).

An online focus group study in Spain, with participants chosen proportionally by gender, age (14–55+), income level, education level, level of Internet use, political opinions, and geographical location, found that three main factors influenced the reported credibility of mis- and disinformation: channel (how an individual knows or discovers the content); source (provenance); and content (including topic and how it is conveyed). Women were more vulnerable to mis- and disinformation than men; older, better-educated, and better-off participants, and those spending less time on the Internet and

identifying as left-wing, were less vulnerable (only the effects of gender and age were statistically significant) (Martínez-Costa et al., 2023).

In the United Kingdom, a study found that one third of Internet users were unaware of the likelihood of finding inaccurate or biased information, and some 30 per cent did know or think about whether the information they encountered was truthful (Ofcom, 2022). Differences have also been found in people's abilities to distinguish between forms of online advertising, and social media platform users are more confident in their abilities than search engine users (Ofcom, 2023a).

Beliefs about the safety of online interactions vary, and people often lack confidence in their ability to address their concerns about the information they engage with online. They may turn to "networks of literacy" (friends, colleagues, and trusted organisations) to navigate online (Carmi & Yates, 2023; Shapiro, 2019). Although online users may be unable to distinguish between news personalisation and platform involvement in targeting them for commercial reasons, they often have concerns about their personal autonomy (Monzer et al., 2020). Although the proportion of young people aged 18–24 who use social media on a weekly basis has been increasing especially for WhatsApp and Instagram in multiple countries (Newman et al., 2023) and also TikTok, research questions whether greater use leads to increased awareness of the risk of harm from online exposure (Global Kids Online, 2019).

Information about children's susceptibility to mis- and disinformation and its impact on their rights and well-being is less well developed than studies of adults (Howard et al., 2021). Children are now impacted by data-intensive technologies throughout their lives, and it is crucial to ensure that their rights are protected (Livingstone et al., 2024; Mahomed et al., 2023). There are difficulties in obtaining information about the changes Big Tech companies make to protect the rights of children, and changes in methods for assuring children's ages have an impact on their rights to freedom of expression and non-discrimination (Wood, 2024). However, research shows that rights-respecting, features designed to respect children's digital rights can contribute to greater enjoyment of children online (Livingstone et al., 2023b). A ten-year study of children in the United Kingdom found that,

children struggled to work out what information they could or could not trust on social media, and many were relatively unmotivated to validate the information they were seeing. Some were keen to show solidarity with views their friends had expressed, without understanding much about the issues under discussion. (Ofcom, 2024a: 7)

In African countries, young people's online engagement has been shown to be associated with relatively low levels of awareness of the impacts of social media use on their rights (Camara et al., 2023). Research in Asia indicates that interpretations of events and online information vary with how people decipher truth in the light of local beliefs (Jayasinghe et al., 2022), while

research in Argentina, Brazil, China, Ghana, India, Jamaica, Russia, South Africa, and the countries in the Caribbean and Southern Africa regions found perceptions of a mix of benefits and risks for online users (Domingos Cordeiro & Cozman, 2024; Dunn et al., 2024).

Understanding automated content moderation

Public understanding about AI and automated content moderation varies, and this is considered a new phase in the public’s growing distrust of digital technologies (Dutta & Lanvin, 2023). Ofcom’s *Online Nation* report found that two thirds of 16–24-year-olds and over half of 25–34-year-olds in the United Kingdom were worried about the impact of generative AI (GenAI) on society (Ofcom, 2023d). An Ipsos survey of some 23,216 adults on public understanding of AI in 30 countries indicated that, overall, a good understanding of what AI is was reported by 67 per cent of participants, but fewer knew what products and services use AI (52%), and 53 per cent said that the use of AI makes them nervous (Ipsos, 2023) (see Table 6.1)

Table 6.1 Understanding AI

	% Agree (30-country average)	Highest %	Lowest %
I have a good understanding of what AI is	67	91 (Indonesia)	41 (Japan)
I know which types of products and services use AI	52	84 (Indonesia)	35 (Belgium) 38 (Japan) 39 (Canada)
Products and services using AI make me nervous	53	67 (Australia)	29 (Japan) 40 (S. Korea & Poland) 43 (Indonesia)
Products and services using AI have more benefits than drawbacks	56	85 (Indonesia)	40 (Canada) 41 (France & Belgium) 42 (United States)

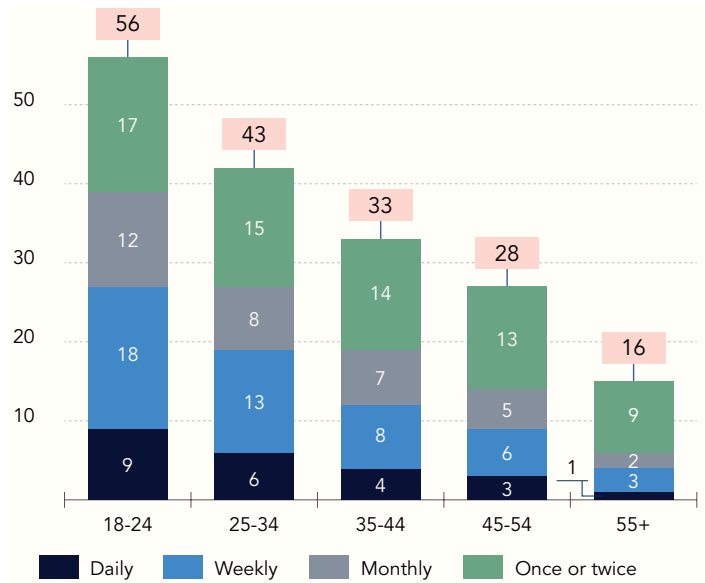
SOURCE: IPSOS, 2025

NOTE: N = 23,216 adults under the age of 75, interviewed 21 March–4 April 2025.

Public attitudes towards the use of AI systems within the news industry vary as well, as demonstrated when people in Mexico, the United Kingdom, and the United States reported being comfortable with using AI tools on a case-by-case basis (generally, most comfortable with GenAI used in news production and least comfortable with tools being used to generate synthetic content). Disclosure about the use of AI systems does not necessarily make news consumers more trusting, and the biggest concern was about being able

to detect mis- and disinformation (Collao, 2024). A survey (approx. 2,000 per country) investigating how often people use AI chatbots or tools in Argentina, Denmark, France, Japan, the United Kingdom, and the United States found that a plurality of people thought GenAI would make their lives better, but a significant majority believed it would worsen their lives. The use of GenAI in news production was expected to make the news less trustworthy and transparent – though more up-to-date – and only a small percentage (8%) thought that the news would be worth paying more attention to if produced by AI tools (Fletcher & Nielsen, 2024). Figure 6.1 shows the proportion of respondents who had ever used ChatGPT.

Figure 6.1 Respondents indicating they have ever used ChatGPT, by age (per cent)



SOURCE: Fletcher & Nielsen, 2024: 12; redesigned by Matthieu Mercier and adapted by Nordicom.

Qualitative research in Brazil suggests that public understanding of how AI-enabled algorithms work influences political disengagement. Brazilian Facebook users have been found to stop engaging politically to avoid an “algorithmic visibility regime” that demeans their civic voices. This was variously attributable to beliefs that algorithms encourage engagement with like-minded users, that online engagement makes citizenship useless, that engagement results in unacceptable sacrifice of values and well-being, and that success in attaining online political visibility does not mean control over one’s visibility (Magalhães, 2018).

A large sample study in the Netherlands, Portugal, and the United States found that algorithmic moderation was believed to be more transparent than

human moderation; ironically, this was particularly so when no explanation was given for the removal of content, and sending users to community guidelines on content removal had negative effects on perceptions of algorithm fairness and trust (Gonçalves et al., 2023). In the United States, predictions of how people respond to algorithm-driven content moderation and selection were found to depend on the heuristic used to explain how this works, that is, whether they consider machines to be more accurate and precise than humans or believe that machines cannot make nuanced subjective judgments. These differences varied by measures of trust, fear, ideology, and ability to use online services (Molina & Sundar, 2022).

The public's awareness of how privacy infringements due to algorithmic-supported data-extraction jeopardise individuals' rights varies widely. Paradoxically, people often claim to be concerned about their privacy, but report being unwilling or unable to take steps to protect it (Nissenbaum, 2011). Research in Europe indicates some public awareness of how political microtargeting infringes on privacy, but uncertainty about who is responsible for data protection, the extent of excessive profiling practices, and the effectiveness of privacy-by-design or by-default features of online services (Blasi Casagran & Vermeulen, 2021). In the United States, online users are shown to be willing to trade company access to their data for "free" access to platform services, even when they report being concerned about unauthorised access to their data (Bright et al., 2022). And greater algorithmic awareness and online skills have been shown to influence online users' privacy concerns and the self-disclosure of data (Shin et al., 2022). In the Philippines, combating online hate speech has been found to require broad social counternarratives and a holistic approach to challenge attacks on gendered, political, and racial identities (Ong & Tapsell, 2022).

Policymaker understanding of AI systems and platforms

Evidence from public hearings and policy investigations suggests that policymakers often do not have a sufficient understanding of how AI systems and platform business models operate (Aula, 2023). Research also indicates that campaigning and lobbying for action against mis- and disinformation can lead to ad hoc responses to the risks (Bunting, 2018; Caplan, 2023) and to a chilling effect on freedom of expression, despite the fact that little is known about the effects of responses on the spread of mis- and disinformation (CERTH et al., 2021). In African countries, policy interventions can restrict freedom of expression or be ineffective (Cunliffe-Jones, 2021).

Some argue that social media gives far-right groups the means to extend their reach: "All of these people are tied together by these loose online networks [...] galvanised by viral online disinformation from unknown and untrustworthy sources" (Tapper, 2024: para. 6, quoting Jacob Davey, Director, Policy and Research, Institute for Strategic Dialogue). In this context,

policymakers often do not have timely evidence to underpin their responses to the spread of false information, and they may not have the power to force a digital platform owner to act. It may also be unclear whether the state should decide what speech is amplified and what is not in the face of high-profile events that generate extensive, inaccurate information.



In the United Kingdom, false claims posted on X that the killer of three young girls in the English town of Southport in July 2024 was a 17-year-old asylum seeker who had arrived in the country on a boat in 2023 were quickly followed by a wave of riots involving far-right groups in several cities and towns. More than 27 million impressions of posts repeating this and similar false claims were recorded (Rusbridger, 2024). Activity on social media platforms revealed “a clear pattern of influencers driving a message for people to gather for protests” (Casciani & BBC Verify, 2024: para. 2). For policymakers, the immediate question was how to respond to the viral spread of false information – the person arrested and charged was not a Muslim, was not a refugee, and was legally residing in the United Kingdom. The Online Safety legislation preventing illegal online speech had not fully come into effect. Government ministers approached the owner of X, Elon Musk, to no effect. There were divided views about whether the X accounts promoting false information should be suppressed or taken down – some said yes, in the interests of security and safety – others, no, in the interests of protecting freedom of expression (Jones, 2024).



Diverse definitions of mis- and disinformation, policies developed in institutional silos, and evidence that is not robust, contradictory, or missing mean that technology, media, and policy professionals have to respond in the absence of evidence that can confirm the workability of their actions (Carson & Wright, 2022; Wasserman, 2022). In Global Majority World countries, as J. Khadijah Abdurahman, founder and director of We Be Imagining at Columbia University’s INCITE Center, put it in an interview (4 March 2024), “policymakers do not understand AI or problems of mis- and disinformation, especially of marginalised or rural groups”. Policymakers within state agencies in these countries may lack understanding about how to tackle mis- and disinformation and gender-related harms and the resources to assess the risks (RIA, 2023b; Van der Spuy, 2023). For example, in the Southern African context, researchers claim that “political leaders lack understanding of what information disorder is and what impact it has”

(Sey et al., 2022: 158). Even if the courts are capable of interpreting laws applicable to platform content governance, “a lack of uniformity, limited capacity and inadequate understanding of AI means that enforcement can be erratic and uneven” (Hlomani, 2023: 2). In addition, actions to address the challenges are “dominated by theoretical paradigms, examples, and case studies drawn from relatively recent experiences in Global North contexts”. The result is that without sufficient conceptual alignment, “people end up producing distinct and incompatible things” (Wasserman, 2022: 7, 112).

Policymakers need to understand digital technologies if they are to make sense of key terms and concepts (Mittelstadt et al., 2023). This applies as much to AI systems as to data and privacy protection and responses to mis- and disinformation. Policy initiatives in response to mis- and disinformation are also criticised for the absence of inclusive participation, failing to encourage community governance, neglecting the interests of marginalised groups, approaching issues in a paternalistic way, or expecting online participants to protect themselves from risks and harm (Schneider, 2022). Policymakers in countries with a high level of press freedom are more likely to pursue holistic approaches to online mis- and disinformation, focusing jointly on election processes and media and education initiatives as well as on measures to combat this information. Countries with greater economic strength are more likely to introduce legislation, with authoritarian countries being more likely to link measures to penal codes (Cipers et al., 2023; Ihlebæk & Sundet, 2023).

Views about what measures should be taken to combat mis- and disinformation differ regionally and by country. For example, in some African countries, public experience of mis- and disinformation and views about rights protection may be influenced by the *ubuntu* philosophy, which recognises that individual rights cannot be fully enjoyed unless the rights of all in a community are respected. This may result in a preference for collective measures aimed at achieving restorative justice instead of relying on government regulation (ADRN, 2024; Bayer et al., 2021; Okyere-Manu, 2023). Research also suggests that an emphasis on universal civic citizenship rights, at the expense of “ethnic conceptions of citizenship”, can downplay “hierarchies of inclusion and exclusion informed by race, ethnicity, class, gender and geography” (Roberts & Bosch, 2023a: 7, citing Nyamnjoh, 2006: 237).

In sub-Saharan Africa, evidence indicates, for instance, that perceptions of the roles of governments, platforms, and users in stopping mis- or disinformation from circulating depend on what people believe about the impacts. Qualitative research (focus groups in six countries) found that mis- and disinformation were perceived as a problem if they had real or perceived negative consequences. Participants acknowledged a personal responsibility for this content, but felt that politicians, political elites, social media platforms, and governments had a greater role to play in combating the problem (Tully et al., 2022; see also Cunliffe-Jones, 2021).

Public support for measures to combat mis- and disinformation varies depending upon the social and political context and culture. Surveys in Mexico, South Korea, the United Kingdom, and the United States found support for government regulatory action, but this support was not evident in Mexico (Chung & Wihbey, 2024). In the United States, online users were found to approve of efforts to increase individual choice “rather than endorsing top-down censorship by platforms or other entities” (Jhaver & Zhang, 2023: 2945). In the United Kingdom, online users reported that they favour more use of fact-checking (Kyriakidou et al., 2023).

Regarding how content is presented and who or what should be responsible for decisions, in Belgium, there was a preference for algorithmic personalisation systems that select content based on similarity, suggesting a desire for means to cope with the volume of online information and an awareness that personalisation can be performed in different ways (Joris et al., 2021). People with a greater understanding of the way algorithmic personalisation systems work are found to have a better understanding of the role of companies and technology developers, and to be influenced by whether they experience online engagement as positive or frustrating (Martens et al., 2023).

In the United States, a study of marginalised youth who depend on social media for their news and political information found that they prefer algorithm-driven online content because they believe this enables them to exercise agency (Kaskazi & Kitzie, 2023). While any intervention in the conduct of online interactions risks undermining freedom of expression (Elsom, 2020; Katsirea, 2018), a narrow margin has been found between those who favour platform use of personal moderation tools (such as downvoting posts, muting, and blocking specific accounts – or, in the case of X (X, 2024), where individuals use Community Notes to challenge accounts whose posts they believe to be false or find offensive) – and those favouring regulatory interventions, with a positive relationship between support for free speech and for platform content moderation (Jhaver & Zhang, 2023; see also Procter et al., 2019; Riedl et al., 2022). These findings might indicate an interest in a “pluralist model of speech regulation [...] in which speech must be regulated in a multi-stakeholder fashion” (Jhaver & Zhang, 2023: 2931, citing Balkin, 2017) and “approval of a shift toward a new approach to content curation that emphasizes individual choice rather than endorsing top-down censorship by platforms or other entities” (Jhaver & Zhang, 2023: 2945). It would be a mistake, however, to generalise results from the United States to other cultures and political systems where the principle of freedom of speech is tempered by the right to be free from harm (Rexhepi, 2023).

A study covering Australia, Brazil, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Spain, the United Kingdom, and the United States, as well as India, Kenya, Nigeria, and South Africa, found that, of those expressing scepticism about the use of algorithms in search engines and social media platforms for selection of news, only 30 per cent agreed that having stories

selected on the basis of their previous consumption was a good way to get news, and 27 per cent indicated they favoured having their news selected by news industry editors (Newman et al., 2023). Research comparing public attitudes to content moderation by algorithms, expert panels, or juries of users has found that expert panels were perceived to have greater legitimacy than juries or algorithms, but the findings were more significant regarding agreement with the decision than they were for the process (Pan et al., 2022). As one study with focus groups in Brazil, India, the United Kingdom, and the United States put it, people's attitudes are shaped by folk theories about partisan political agendas and commercial considerations (Mont'Alverne et al., 2023).

A study on Brazil, India, and Germany found that interviewees (academics, civil society, policymakers, and industry representatives) agreed that mis- and disinformation "brings to the surface and magnifies the tensions between national jurisdictions and a technology which is transnational in nature", that regulation is "usually toothless when problems originate beyond national borders", and that regulation may be exploited by authoritarian governments and curb competition (Marda & Milan, 2018: 10, 11). "Different democracies with distinctive underlying sociopolitical contexts and cultures have varying ideas and translate values in nuanced ways regarding platform regulation", as evidenced by survey research in Mexico, South Korea, the United Kingdom, and the United States (Chung & Wihbey, 2024: 4546).

Investigations in Asia, the Caribbean, Latin America, the Middle East, North Africa, and sub-Saharan Africa recommend a "critical praxis" response to mis- and disinformation linking scholarship with practical action so that "contextual imperatives, lived experience and local knowledge feed into policy measures" (Wasserman, 2022: 21). With people's awareness of and ability to navigate information ecosystems varying by country, gender, race, class, culture, socioeconomic status, and more, it is unsurprising that what the public would like states, the private sector, and other organisations to do to mitigate harms linked to mis- and disinformation depends on their everyday experience of how mis- and disinformation manifests. This requires ethnographic research, especially with unconnected or poorly connected communities that are often excluded (Wasserman, 2022). In addition, attitudes about strategies towards building trust in news media (based on an international survey, $N = 8,229$) depend on social, political, and cultural cleavages that are "unique to some degree to each country. Between the Global North and Global South [...] different demographic segments have distinctly different relationships with civic institutions, which shapes people's expectations towards news media as an institution as well" and there is "no silver bullet for cultivating the public's trust" (Banerjee et al., 2023: 68).

Evidence on the public's appetite for policy interventions to counter illegal and harmful mis- and disinformation suggest a slight preference for content moderation undertaken by online participants over content moderation by platforms or state regulation, but the evidence is not conclusive, or inclusive of

all countries. Westerncentrism is present in many academic papers and reports (Nguyễn et al., 2022), and the views of regional organisations in the Global Majority World must be taken into account (Chairpersons Organizations, 2023; UNESCO, 2023c). In relation to the news media, epistemologies that reduce “truth” to simple definitions of objectivity or impartiality need to be avoided, alongside efforts to mobilise lobbying to insist that tech companies do more to tackle mis- and disinformation in a transparent way in order to minimise unevidenced responses by Big Tech companies and by states, especially during crises, including war and conflict (Gagliardone et al., 2015; Haas & Kettemann, 2024). Multistakeholder dialogue is essential with attention to whether marginalised and vulnerable populations are consulted in ways that are informed by principles, including feminist principles, that encourage self-determination (Marda & Milan, 2018; IT for Change, 2023a, b).

In summary, international agreements may be clear that government measures “to suppress dissent and to control public communication” must be “necessary and proportionate” to protect legitimate interests under international law (OSCE, 2017: 2). Nevertheless, as Kaye (2015), former United Nations special rapporteur to the Human Rights Council, has observed, the potential for censorship is always present when mis- and disinformation or hate speech circulates. Governance practices could become more consistent with human rights standards if policymakers ask themselves whether legislation or nongovernmental methods are the best approach to counter the harms of mis- and disinformation. How mis- and disinformation are defined, independent oversight of interpretations and adjudication processes, and whether content governance should be discretionary or mandatory, are all issues that must be addressed if governance is to support inclusive and just information ecosystems (CNTI, 2024; Matasick et al., 2020; Pradhan, 2021).

Measures to combat mis- and disinformation

Efforts to inform the public about the authenticity and credibility of information play a major role in combatting mis- and disinformation. There is a large mix of such efforts and many studies evaluating their effectiveness (Mendoza et al., 2023). These generally emphasise that combatting mis- and disinformation requires an effort that transcends individual or civil society stakeholders to include actions by governments and the owners of digital platforms (Berger et al., 2023a).

Table 6.2 gives an overview of common mis- and disinformation countermeasures, indicating what is known about them, whether they have been found to be effective, and whether they scale. Actions include improving public information, government action, and platform action. This review of countermeasures concludes that “there is no silver bullet or ‘best’ policy option” (Bateman & Jackson, 2024: 2) and notes that much of the evidence is drawn from work by United States-based scholars. Voelkel and colleagues

(2024), for example, found in the United States that 23 of 25 treatments aimed at reducing antidemocratic attitudes and partisan animosity were effective to some degree.

Table 6.2 Selected mis- and disinformation countermeasures

Type of Action	Intervention	How much is known?	How effective does it seem?	How easily does it scale?
Public information	Supporting local journalism	Modest	Significant	Difficult
	Media literacy education	Significant	Significant	Difficult
	Fact-checking	Significant	Modest	Modest
	Labelling social media content	Modest	Modest	Easy
	Counter-messaging strategies	Modest	Modest	Difficult
Government action	Cybersecurity for elections and campaigns	Modest	Modest	Modest
	Statecraft deterrence and disruption	Modest	Limited	Modest
Platform action	Removing inauthentic asset networks	Limited	Modest	Modest
	Reducing data collection and targeted ads	Modest	Limited	Difficult
	Changing recommendation algorithms	Limited	Significant	Modest

SOURCE: Bateman & Jackson, 2024: 5

Where measures such as debunking, accuracy prompts, and media literacy tips are tested globally, their effectiveness is found to be sensitive to cultural contexts. GenAI tools have become widely available at relatively low cost, require little or no technical expertise, and their use to create political mis- and disinformation is growing. Some tools for detecting this information are found to do so with 80 to 90 per cent accuracy on content created in the Global North, but they are much less effective on content created in the Global Majority World. This is due to biases in their training data and the fact that as the “tools were developed, they were prioritized for particular markets”, privileging “English language – US-accented English – or faces predominant in the Western world” (Elliott, 2024: para. 4, quoting S. Gregory, Witness.org).

Public information and fact-checking

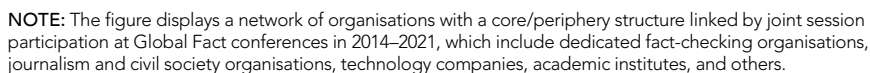
Countermeasures to provide public information about mis- and disinformation, and specifically fact-checking, are becoming widespread, with numerous initiatives funded in some cases by media organisations, in

others by governments or independent organisations (see Table 6.2). Research is devoted to developing AI tools to identify mis- and disinformation, and several of these tools are used by professionals in their fact-checking work. However, these tools on their own cannot counter the threat of mis- and disinformation (EC, 2024b; Marinova et al., 2020; Teyssou et al., 2017).

In this section, we focus on fact-checking defined as “the practice of systematically publishing assessments of the validity of claims made by public officials and institutions with an explicit attempt to identify whether a claim is factual” (Walter et al., 2020: 351). In principle, fact-checking is like expert panel content moderation, but it maintains user agency by leaving the user to determine its relevance. Fact-checking by nongovernmental organisations emerged in the early 2000s. Independent fact-checking organisations numbered around 113 by 2016 worldwide (Graves & Cherubini, 2016), growing rapidly in response to the increasing volume of mis- and disinformation to an estimated 447 active organisations as of August 2025 (Duke Reporters’ Lab, 2025) (see Figure 6.2).

Meta was paying fact-checkers in 119 countries to identify misinformation on Facebook and Instagram in 2024 (Lauer & Graves, 2024). After the company said it would stop fact-checking in the United States in early 2025, it is unclear how this policy will affect other countries. In May 2025, the Canadian company, Telus International, which had been contracted by Meta to moderate content, announced it was cutting 2,000 fact-checking jobs in Spain (Bontemps, 2025), and the phrase “information integrity” was being characterised as a politically neutral term compared with mis- and disinformation, which Big Tech United States-based companies, the Trump Administration, and other governments had started to treat as being politically charged (Full Fact, 2025).

SOURCE: Lauer & Graves, 2024: 3513; adapted by Nordicom.



As fact-checking organisations struggle with the growing volume of “checkworthy” claims on social media, there is increasing reliance on machine learning techniques (Quelle & Bovet, 2024). Many fact-checking organisations are committed to retaining a “human-in-the-loop” (Procter et al., 2023), but automated fact-checking may be inevitable in the future. Advances in AI systems, including GenAI tools, have succeeded in lowering the costs of mis- and disinformation production, while making it accessible to more people with little technical knowhow or training (Xu, 2023). Fact-checking is being automated using natural language processing, machine learning, knowledge representation, and a variety of databases (Guo et al., 2022). A study examining whether large language models (LLMs) might also be used to counter these kinds of information demonstrated that, when used with prompts specifically designed for the purpose, LLMs may be a useful tool for detecting LLM-generated mis- and disinformation, but these tools struggle to consistently identify this information (Ghosal et al., 2023; Jiang et al., 2024). Researchers are also investigating the use of these tools to detect and mitigate the effects of hate speech, including extending the language-detection capabilities and detecting implicit or coded speech and testing whether these tools result in reductions of user engagement (for a review of studies, see Albladi et al., 2025; see also Podolak et al., 2024).

The applications of these models depend upon the motivations of their developers, and fact-checking is not a static process where effectiveness can be assessed without reference to changes in techniques and practices over time or changes in how people engage with fact-checked information. A study during the 2020 midterm elections in the United States assessing the effectiveness of interventions to combat mis- and disinformation on X/Twitter concluded that fact-checking by professional fact-checkers was less effective than faster algorithmic fact-checking, even though the latter was more prone to error (Guriev et al., 2023). As these tools to combat mis- and disinformation are developed, an AI systems “arms race” is making mis- and disinformation less detectable (Innes et al., 2023), and evidence from studies on the effectiveness of fact-checking is mixed.

The effectiveness of fact-checking has been evaluated using quantitative experimental designs and qualitative methods involving focus groups, interviews, and surveys. A study involving a total of 28 experiments in which participants in Argentina, Nigeria, South Africa, and the United Kingdom evaluated 22 fact-checks on topics including Covid-19, local politics, and crime found that fact-checks reduced beliefs in mis- and disinformation, but there was also evidence of a so-called backfire effect where the opposite was found (Porter & Wood, 2021). An experiment in Australia investigating public trust in third-party fact-checkers of political claims (some 1,600 participants) found, contrary to expectations, that third-party fact-checking had a negative influence on trust in a fact-checked news story – indicating the presence of this backfire effect on media trust. However, the public’s trust in news stories and

fact-checks done on them depended on trust in the media, indicating that if third-party fact-checks are intended to help restore trust in media, they need to be conducted in different ways (Carson & Gibbons, 2023).

A review of empirical evidence of the efficacy of fact-checking (in Brazil, Malaysia, Turkey, the United Kingdom, and the United States) concluded that fact-checking efforts to hold politicians to account were not meeting the expectations of policymakers (Vinhas & Bastos, 2022). The transparency of sources, funds, and methods of fact-checking is key to whether the public trusts its use (Brandtzaeg et al., 2018). Fact-checking transparency was compared in a study with fact-checkers in Australia, Brazil, India, South Africa, the United Kingdom, and the United States, finding that South African fact-checking undertaken by nongovernmental organisations scored higher in transparency than when undertaken by newsrooms and other fact-checkers; and India scored higher than the United States and the United Kingdom (Ye, 2023). Some evidence on the impact of fact-checking suggests that it has a limited effect on responses to mis- and disinformation. Two qualitative surveys ($N = 1,065$; $N = 542$) in the United Kingdom focusing on the understanding of, and engagement with, fact-checking found that it had a minimal impact on news consumption and that participants had a desire for more television news fact-checking, although the perceived credibility of fact-checking tended to reflect people's beliefs in the impartiality (or otherwise) of news channels (Kyriakidou et al., 2023).

Fact-checking during elections includes detecting, tracking, and countering mis- and disinformation and content curation, emphasising community standards and algorithm changes, and sometimes demonetising content by requiring platforms to stop those who generate clickbait or counterfeit news from receiving an income. These approaches include emphasising ethical practice, media and information literacy, and empowering individuals and journalists to use content verification tools and web content indicators (Cipers et al., 2023). A meta-analysis examining the effectiveness of fact-checking in correcting political mis- and disinformation indicated that it can have a positive and significant effect by correcting beliefs, positively affecting beliefs “irrespective of political ideology, pre-existing positions, context (campaign vs. routine) and whether it refutes the entire false statement or just parts of a statement” (Walter et al., 2020: 366). The effects of fact-checking on beliefs in this study were not very strong, however, with effects becoming almost negligible in cases where a study design mimics real-life scenarios. Thus, fact-checking may be more effective for strengthening preexisting beliefs than for countering mis- and disinformation. A systematic review of experimental studies across 40 countries of people's accuracy ratings of true and false fact-checked news found that people are more likely to rate true news as accurate than fact-checked false news, but that they disbelieve true news slightly more than they believe fact-checked false news. This study argued for interventions that support openness to opposing viewpoints (Pfänder & Altay, 2025). In

the United States, perceptions of the objectivity of fact-checking organisations are influenced by political and ideological divides, reducing the perceived trustworthiness of fact-checks. When fact-checking is most needed during electoral campaigns, audiences may give little attention to fact-checks that contradict their political views and preferences (Walter et al., 2020).

Fact-checking is an important contribution to broader efforts to hold power to account. It can create a reverberating “trail of accountability” of politicians by tarnishing reputations when they promote divisive issues. In addition to fact-checking, other measures include indicating which news outlets adhere to international journalism standards by using labels, using tools to prioritise content that is deemed to be trustworthy, and promoting this trustworthy news to audiences and advertisers. Initiatives by NewsGuard and The Trust Project (both nonprofits in the United States), the Publicis Groupe (France), and the Journalism Trust Initiative (Reporters Without Borders), for example, contribute in these ways (Mazzoli, 2023).

Research findings on the effectiveness of fact-checking cannot be generalised across countries since contexts vary. In addition, while fact-checking may detect problems with a particular news story, which would justify a decline in news media trust, this may not be linked to a decline in trust in the news outlet as a whole, and the results of research in this area do not identify any single approach that can be relied upon to enhance people’s resilience to mis- and disinformation. The next section looks at legislative measures and codes of practice introduced through government actions.

Government action: Legislation and codes of practice

Government legislation and codes of practice are introduced to deter or disrupt mis- and disinformation and its circulation, especially during elections (see Table 6.2). Although laws have been in place against the communication of specific types of false information during elections, the increasing possibilities for actors to covertly disseminate misleading information has led many countries and regions to update their regulations, particularly with a focus on foreign interference in elections (e.g., France and Australia). Such measures may aim to increase the transparency of political online advertising and to ensure the availability of publicly accessible ad archives, but Big Tech compliance with these regulations varies (Tenove, 2020). Struggles between governments and Big Tech companies highlight the tension between freedom of speech and freedom from harm and the challenges of enforcing policies that respect national sovereignty and domestic law in the face of the foreign-owned digital platforms’ power (Curzi de Mendonça, 2024).



A Brazilian judge issued a banning order on X after Elon Musk refused an earlier court order to suspend more than 140 accounts, which the same judge had determined were a threat to Brazilian democracy. The judge declared that Musk was allowing “the massive spread of disinformation, hate speech and attacks on the democratic rule of law, violating the free choice of the electorate, by keeping voters away from real and accurate information” (Nicas & Conger, 2024: para. 9). Musk responded that the judge was an “evil dictator” (Shapero, 2024: para. 1). X was later reported to have circumvented the ban by an update to its network that enabled alternative routes for Brazilian users to access X (Murphy et al., 2024). In other cases, X has complied with government demands to ban accounts and content, for example, when it demoted or banned content that was critical of India’s Prime Minister Narendra Modi during farmer agricultural policy protests in 2024 (Sankaran, 2024).



Of laws responding to false information from 2016 to 2020 in African countries, some have passed or amended laws that penalise the publication or broadcast of what the authorities consider to be false, or they draw upon other laws using arbitrary definitions of false content with an unwelcome effect on freedom of expression – and also failing to limit the spread of mis- and disinformation (in Benin, Burkina Faso, Côte d’Ivoire, Ethiopia, Kenya, Malawi, Niger, Nigeria, Senegal, South Africa, and Uganda) (Cunliffe-Jones, 2021). Government use of digital technologies as tools for surveillance, control, and censorship deters political expression and engagement, and research points to the increasing involvement of the military, police, state information services, and security consultants (Casero-Ripollés et al., 2023).

Different approaches are present in democratic and authoritarian regimes. A study of government responses to online mis- and disinformation in 103 countries and by 10 international and regional organisations across six continents, found that authoritarian regimes prefer sufficiently vague responses to allow them to repress opposition and criticism (e.g., a cybercrime law in Zimbabwe, enforcement of Article 80 of the Constitution in Tunisia, and cybercrime laws in Nicaragua). The less democratic the country, the less likely they are to implement election-specific responses to maintain the integrity of the electoral process, and repressive legislation is not limited to authoritarian states, since they are present in democracies such as India and Mexico (Cipers et al., 2023).

The European Union’s *Code of Practice on Disinformation*, which was established to oversee content governance, has been taken up beyond the region, and most platforms employ some form of content moderation or oversight system (EC, 2022e; see also Chapter 5). Monitoring studies indicate that the

code is a useful tool, but there is room for improvement (EC, 2020a), and the transparency of content removal policies and practices is criticised for being too weak (Solomon et al., 2021). The code has been criticised for being too general; a lack of compliance and enforcement tools; the absence of a shared terminology; and a lack of granular platform data for independent assessment. Other criticisms include the risk of the privatisation of censorship (i.e., delegation of important content-based choices to platforms); insufficient collaboration; unclear or unstated reporting on regional actions; and a lack of standard procedures to verify implementation (ERGA, 2020; Monti, 2020; Nenadić et al., 2023; Pamment, 2020). Based on digital platform annual reports, others have concluded that as a form of soft governance, the *Code of Practice on Disinformation* is effective under certain conditions due to its flexible reporting requirements and to encouraging the platforms to give a higher priority to combating mis- and disinformation; however, this voluntary code is also seen as “soft governance [which] permits companies to window dress low compliance by reporting cost effective operational changes” (Borz et al., 2024: 18).

Investment in mechanisms to combat mis- and disinformation varies across regions. It is concentrated in the Global North, which sees more consistent enforcement of the platforms’ own rules. In smaller markets and content in languages other than English, these mechanisms receive much less attention (Dergacheva et al., 2023; Global Witness, 2023; UN, 2023a). Systems for automatic content moderation are largely trained on English-language content because of the gap between the quantity, quality, and diversity of LLM training data in English and other languages (Nicholas & Bhatia, 2023). A report on the effectiveness of a broad range of countermeasures found that fewer than one-fifth of studies in a systematic review of academic studies reported effective countermeasures, that evaluations depended on the type of intervention and information, and there was uncertainty about the size of effects (IPIE, 2023a; see also IPIE, 2023, b, c). A study of institutional and regulatory frameworks in Central and West Africa highlighted the absence of effective policies for fighting mis- and disinformation, with 46 per cent of women and girls reporting that false narratives are used to dissuade women from participating in the public sphere (Zibi Fama, 2024).

Efforts to combat mis- and disinformation are also limited by the economic fragility of news media organisations and their susceptibility to capture by governments. Research emphasises the need for legal safeguards to prevent the risk of political influence (Radsch, 2023d). Negative impacts on freedom of expression can occur even when attempts to govern in good faith are made. For example, in a move that was widely criticised, Italy’s “Operating Protocol for the Fight Against the Diffusion of Fake News through the Web on the Occasion of the Election Campaign for the 2018 Political Elections” gave the Postal and Communications Police [Polizia Postale e delle Comunicazioni] the authority to determine whether online claims were false or biased and to recommend judicial action (Kaye, 2018).

Platform action to address mis- and disinformation

The Big Tech companies have responded to mounting public scrutiny by implementing measures – some temporary and linked to circumstances such as Covid-19 or specific electoral campaigns. As indicated in Table 6.2, the Big Tech companies can remove inauthentic asset networks (i.e., detect and remove accounts that misrepresent who social media influence operations represent), reduce data collection and targeted ads, and change their recommendation algorithms. Platform actions can additionally include flagging and reviewing content; filtering, limiting, blocking, or removing pieces of content; promoting or demoting content; increasing the transparency of sponsored content; and implementing user involvement and appeal mechanisms (Bontcheva et al., 2020). The platform owners are generally reluctant to censor content or permanently ban prominent users (i.e., political figures) who promote mis- or disinformation, although they have been found to do so when they become embroiled in politically partisan disputes.

A voluntary AI Elections Accord to take reasonable precautions to prevent AI tools from being used to generate inaccurate or false information that could contribute to the disruption of elections being held in 2024 was signed by Adobe, Anthropic, ElevenLabs, GitHub, Google, LG AI Research, LinkedIn, McAfee, Meta, Microsoft, Nota, OpenAI, TikTok, and Truepic (AI Elections Accord, 2024; O'Brien & Swenson, 2024). The text acknowledged that “deceptive AI Election content can deceive the public in ways that jeopardise the integrity of electoral processes”, and the companies called for a “whole of society approach”, noting that no individual solution, or even combination of solutions, would “fully mitigate the risks” and there was a call for public education to mitigate risks (AI Elections Accord, 2024: 1). The accord was in place for one year, and there are calls for renewed agreements and collaborations in the face of a growing use of deepfake videos and a proliferation of AI-generated news sites. Civil society organisations have called for formal platform commitments to the *Santa Clara Principles 2.0*, updated in 2022 and endorsed by Apple, Meta, Google, Reddit, X, and Github at the time, which address the transparency and safeguards for misleading user-generated content (Santa Clara Principles.org, 2022; see also Palladino et al., 2025 for initiatives by other civil society organisations).

When platforms such as Facebook, YouTube, and X use their algorithms to amplify engagement with content to promote growth, communities of colour, women, religious minorities, and members of the LGBTQ+ community can be ignored when harms associated with content that targets them are unaddressed (Díaz & Hecht-Felella, 2021). When Big Tech companies do moderate what they determine to be unacceptable content, a preferred approach is to act against suspicious behaviour on user accounts by deprioritising the visibility of content in news feeds using a mix of automated and human curation content moderation practices (Saurwein & Spencer-Smith, 2020).

Measures to combat mis- and disinformation using algorithms that amplify content engagement are weakened when the Big Tech companies lay off staff, weaken privacy policies, or impose limits on fact-checking (Kennis, 2024). Meta, YouTube (Google), and X, for instance, have been criticised for engaging in a new form of “private multilateralism” when they appear to comply with legislation, but go beyond requirements to mitigate reputational damage or to ward off further legislation, with the result that they have too much discretion over who is silenced (Borelli, 2023). When a platform bans a user’s content from appearing without notification, this stifles the user’s free speech, and this receives less attention in comparison to harms associated with the virality of content (Horten, 2023). The Big Tech companies’ use of content detection and takedown tools are also criticised for tackling specific forms of mis- and disinformation or propaganda, for example, ISIS- or al Qaeda-generated content, with findings that these measures do not work well in other contexts (Mitts et al., 2022).

Simple solutions to the harms of mis- and disinformation, such as insisting on non-discrimination, are not viable, and “there is no such thing as a ‘neutral’ recommendation algorithm” (Llansó et al., 2020: 3) (see Chapter 3). Approaches to mitigating the biases of algorithmic personalisation systems include offering users ranking options in their feeds to enhance their agency, supporting media and information literacy, and boosting transparency around the design and performance of these systems. Since various forms of content moderation do not provide disincentives for the production of mis- and disinformation, it is suggested that algorithm design changes could alter content rankings to reward productive interactions, place limits on the dissemination of content, and “support the on-platform work of peacebuilders”, so that everyone is able to thrive (Stray et al., 2023: 6).

The long-term effects of applying countermeasures have been tested with those who are exposed to mis- and disinformation, with some evidence that effectiveness decreases with time. Comparative research is limited by significant variability in methods (e.g., test stimuli using news headlines, real-world claims, or websites) and ways of measuring outcomes (e.g., belief or credibility ratings, behavioural measures) (Kozyreva et al., 2024). Research is also limited by data availability. Progress is slow by companies (e.g., OpenAI, Google, Anthropic, Inflection, Meta, Midjourney, and Cohere) towards meeting commitments to make data available (e.g., by providing public application programming interfaces) and deep access to data (e.g., LLM base models, model families, and filters or content moderation systems). Harrington and Vermeulen (2024: 35) concluded that AI model developers had retained exclusive control “over the majority of research access initiatives”. Reluctance to provide data access in a way that would help to increase the transparency of Big Tech platform operations is illustrated by Meta’s CrowdTangle.



CrowdTangle, purchased by Meta in 2016, was valued for its use in mis- and disinformation monitoring. The tool allowed researchers, journalists, and fact-checkers to explore public content posted on social media platforms, including Facebook, X/Twitter, Instagram, and Reddit. In 2024, Meta shut the tool down, explaining that this was due to the European Union's Digital Services Act. The Meta Content Library was said to provide the same services that were previously available through CrowdTangle; however, usage limits meant that many CrowdTangle users were not allowed access. This was criticised as a political move to censor partisan information close to the United States presidential election. It is also claimed that the Meta Content Library has only one per cent of CrowdTangle's features, and the library has been criticised by the European Commission for reducing platform transparency and access to data (Bellan, 2024; Gotfredsen & Dowling, 2024; Kroet, 2024).



Reluctance on the part of Big Tech companies to make their operations more transparent has led to proposals for achieving “meaningful” platform transparency by mandating interoperable platform transparency. An Internet Governance Forum proposal calls for platform reports that provide quantitative and qualitative contextual information when they make data available for research purposes, sharing detailed and intelligible information in standardised formats, and making data continuously available in an interoperable, understandable, and machine-readable format that is auditable by third parties (Belli et al., 2022).

Achieving accountability is challenging, because most of the largest companies investing in data-intensive products and services are based in the United States and China. And in most countries, the powers of policymakers to force these Big Tech companies to change how they operate are relatively weak. As the Big Tech companies have become the “new governors” with “little direct accountability to their users” (Klonick, 2017: 1), and as the use of AI systems is more extensive, the ethical use of these technologies has been judged to be variable in research covering 75 countries, with indications of reasons for both optimism and concern, especially in relation to whether human rights protections (mainly privacy and freedom of expression) and democratic values are being upheld (CAIDP, 2025).

In summary, diverse government, Big Tech, and civil society approaches to combating mis- and disinformation, including content moderation, are justified by diverse contexts. However, the evidence indicates weaknesses in enforcement, transparency, and opportunities to appeal government

and platform moderation decisions across countries. It is also important to recognise that some politically motivated mis- or disinformation is satire aimed at ridiculing elites and, in African countries in which trust in mainstream news media is often low, this type of information has long been present before mis- and disinformation became a preoccupation in the Global North (Morales et al., 2021). In the case of all these measures to govern the production and circulation within information ecosystems, it is crucial to take values into account, including solidarity and interconnectedness, in order to explain why people are motivated to produce and share mis- or disinformation and to assess whether responses are just and respect human rights.

Alternative data governance practices

Concerns about the outcomes of content governance and the way people are being surveilled in the process of collecting their data, especially when aided by AI systems, can manifest in resistance to the data collection practices of the Big Tech companies, and to approaches to combating mis- and disinformation when they impact negatively on individuals' human rights and lead to collective population harms (see Chapter 5). Efforts may be made to game an AI system by entering fake data inputs or to oppose the algorithmic content management practices of Big Tech companies. Achieving just and inclusive information ecosystems depends not just on the legislation and regulations that are put in place but on the agency of individuals and groups to resist the power of technology companies and governments when the outcomes of their initiatives are untransparent and result in risks and harms that persist.

Epistemic privileging and data justice

Resistance aimed at achieving data justice within information ecosystems is occurring through efforts to devise alternative data governance frameworks and practices. Resistance requires an understanding of epistemic privileging; that is, how knowledge, structures, and practices combine with technological affordances to shape what counts as “good” knowledge and whose knowledge matters in decisions about the governance of information ecosystems (Wihbey, 2024; Wu, 2017). Current data governance arrangements have an impact on the quality of information that is produced and circulated, with multiple risks and realities of reproducing or worsening socioeconomic, gender, racial, and other forms of discrimination, and a growing literature examines how this occurs (see Casati, 2013; Couldry & Mejias, 2019; Fuchs, 2023; Hintz et al., 2019; Mejias & Couldry, 2024; Papacharissi, 2013; Powell, 2021; Söderström & Datta, 2023; Vaidhyanathan, 2011, 2022; van Dijck et al., 2018a).

Epistemic privileging by class (caste), race, gender, political affiliation, and economic status is not a new phenomenon, because information is always interpreted in the light of power structures (Horowitz et al., 2024) and the epistemic privileging of someone's knowledge is inevitable – the issue is whose knowledge and with what consequences (Rouvroy & Berns, 2013). Biases and exclusions that occur in the processes of technological innovation, and with the application of these technologies in the Global North and Global Majority World, need to be exposed if justice is to be achieved, and this means that governance – institutions and practices – must change when they are found to be marginalising and harming populations (Chambers, 1997; Thakur & Madrigal, 2022; Willems, 2014b). Inequities cannot be addressed merely by balancing priorities for investment in digital technology with the public interest, for instance, by prohibiting specific digital applications while promoting massive investment with little regard for safety and potentials for discrimination. A profoundly deeper analysis is needed of how historical and contemporary power dynamics shape information ecosystems in ways that reinforce entrenched inequities and influence which voices are heard, and which ones are silenced.

Power dynamics that give rise to epistemic privileging mean that no configuration of technological affordances is universally beneficial. The structures and processes of knowledge production and circulation within information ecosystems that give rise to these affordances contribute to the reinforcement of existing hierarchies in society (Foucault, 1980), perpetuating disparities in both illiberal and liberal governance regimes (Sodré, 2021). In the case of platform datafication processes that lead to our increased reliance on data, perverse economic incentives are created whereby marketeers operate to capture people's attention in a data economy that thrives on data extraction and monetisation (Misra, 2022). In the current configuration of data capitalism, the result is a “distribution of power that is asymmetrical and weighted toward actors who have access and the capability to make sense of data” (West, 2019: 23). When research attends to how epistemic privileging works, this encourages a departure from the assumptions that data universalism is to be valued and that the experience of datafication is homogeneous. This research reveals how the extraction, accumulation, and commodification of data differentially influences people's lives. It also provides insight into how power dynamics leading to injustices can be resisted by supporting counter-epistemic processes and alternative data governance institutions and practices (Arriagada et al., 2023; Cieslik & Margócsy, 2022; Horst et al., 2024; Milan & Treré, 2019; Segura & Waisbord, 2019).

Data activism that aims to resist exploitative data practices must contend with a “double helix” whereby data is extracted from places and people in the Global Majority World by the technology industry, and knowledge about this process is extracted by researchers and their institutions in the Global North. Researchers in the Global North (or those in positions of power

in the Global Majority World) need to be sensitised to these material and epistemic processes if data governance is to embrace the needs of those who are marginalised or discriminated against (Engel & Noske-Turner, 2018; Lehoué, 2022).

Overcoming information access divides and differences in capacities to interpret online information by addressing digital divides, including the emerging AI divide, is insufficient (see Chapter 1) if data governance arrangements are to become more equitable. The social, cultural, and political institutions that hinder meaningful and beneficial uses of technology, and the barriers to the generation of alternative knowledge about how to govern these technologies, must also be understood. Although algorithmic bias has been a feature of computational systems for decades (see Chapters 3 and 4), its recent manifestation creates new demonstrations of epistemic privileging that enable social media, search engines, and a variety of AI applications to reinforce racial and gender stereotypes and privilege – often Western – knowledge (Noble, 2018). This is not an incidental byproduct of progressive innovation and adoption of digital technologies; it is a reflection of the sociopolitical contexts within which choices are made about how these technologies are developed and deployed.

What is needed is a reimagining of information ecosystems with the potential to privilege fairness, justice, and human rights, not just in principle, but in practice. This requires a decolonisation of knowledge to enable the experience of those in the Global Majority World to inform choices about how digital technologies are developed and governed. Research contextualised by experience that does not reveal sources of exclusion or inequitable inclusion within information ecosystems is common (Schoon et al., 2020). However, there are many examples of communities that are organising to resist epistemic privileging that results in the oppression of individuals and groups as a result of Big Tech-driven datafication strategies. These examples involve political engagement aimed at achieving data justice in parallel to, in resistance against, or separately from, government- or corporate-led data governance (see Alaimo & Kallinikos, 2024; Gillwald et al., 2022; Gurumurthy & Chami, 2024; Milan et al., 2021; Niklas & Dencik, 2024; Santos & Ndlovu, 2022).

Resistance to corporate datafication strategies

Alternative data governance principles and practices are being developed that are responsive to the needs of specific locales. Ways of creating and sharing knowledge and information (digital or otherwise) that do not depend on digital products owned by Big Tech companies are being developed in countries and regions as diverse as Argentina, Brazil, the Caribbean, China, Ghana, India, Jamaica, the Philippines, Russia, South Africa and Southern Africa, and the United Kingdom. These initiatives have in common

efforts to resist algorithmic authority imposed by Big Tech companies by reimagining alternatives to algorithmic injustice. They also abandon a cybernetic ideology that couples technology innovation with modernity and a narrow view of progress that assumes there is only one direction of technological change (see Bhat, 2021; Buolamwini, 2023; Cammaerts & Mansell, 2020; Caballero & Monje, 2024; Carroll et al., 2019; Domingos Cordeiro & Cozman, 2024; Dutta & Pal, 2020; Mejias & Couldry, 2024; Mansell, 2012; Noble, 2018; Roberts & Ali, 2021; The Tierra Común Network, 2023).

Resistance strategies aiming to devise alternative data governance approaches do not necessarily aspire to be scalable to mirror the Big Tech companies, although they may seek to do so. They do aim to consult marginalised and vulnerable populations to devise data management solutions based on their judgments and to enable participatory action that leads to digital self-determination (Medrado & Verdegem, 2024; Zhang et al., 2023). They offer models for how communities can engage in “deepening democracy” through participatory governance (Fung & Wright, 2003), and they provide insight into what digital democracy and data justice could look like (Ford, 2019).

Central to many of these initiatives is an emphasis on the “collectivity” – that is, on issues that have societal implications instead of focusing principally on individual rights protections (Dencik, 2025). Research on active citizenship and democratic participation in the Global North has shown that resistance strategies do not materialise in a vacuum and must take account of specific characteristics of institutional regimes (Hintz et al., 2019; Isin & Ruppert, 2020). Similarly, in the Global Majority World, resistance to data extraction requires an understanding of how individual and collective action can be positioned to ensure that choices about the development and use of digital services are not skewed by the preferences of distant Big Tech companies or autocratic states (Roberts & Bosch, 2023b). While “one cannot simply brush away these new forms of dispossession and inequality with a single new law, a revolutionary technology or even a social revolution” (Mejias & Couldry, 2024: 206), empowering approaches to data governance can emerge through a variety of strategies and tactics that aim to resist data extractivism and develop governance frameworks and practices, even on a contested battleground (see Table 6.3).

Table 6.3 Corporate datafication resistance strategies and tactics

Actor	Tactic	Purpose
Individual self-defence strategies	Adopt privacy-enhancing technologies such as virtual private networks (VPNs) and encrypted messaging applications	Digital self-defence practices and digital dissent
	Opt out of dominant social media platforms	
	Remove personal information from public and private datasets	
Public interest alternative news media	Investigate corporate data practices and concentrations of corporate power, with a focus on the impacts of practices on marginalised groups	Enhance public awareness of the harms of poorly constrained commercial datafication
		Hold technology companies accountable to limit expansion of corporate power
Community collaborative strategies – Indigenous communities and municipal initiatives	Produce or collect data relevant to the needs of communities	Establish a citizen-first, rather than technology-first, approach to data governance
	Develop community-owned and -run platforms for recording and sharing information	
	Establish community principles for data with or about communities and municipalities	Create alternative data norms. Model the creation of new data norms for other communities
	Demand that policy around data responds to the needs of municipalities	Democratise legislative processes around datafication, and fight for new ordinances
Social entrepreneurs and community-controlled technologies and data practices	Partner with local communities to develop new technologies based on non-extractive data practices	Provide alternatives for individuals and communities to avoid contracting with multinational companies
	Develop open-source software applications	Provide personalised tech support and digital literacy training
	Develop public datasets with local data capturing local knowledge	
New national-level decentralised data governance frameworks	Build networks and own and control data	Encourage community data hubs, decentralised data infrastructures, local data analytics, data lockers and cooperatives, and public data infrastructures
	Develop new data governance frameworks that preserve local control of data	
	Develop commons-based data governance	

Continued

Table 6.3 Continued

Actor	Tactic	Purpose
Civil society organisations, researchers, and philanthropic organisations	Establish systems of social support for communities harmed by tech dominance and datafication	Mitigate harms to individuals and communities, and remove barriers to organising and other forms of democratic participation for those most negatively impacted by datafication
	Engage with local, national, and international human rights bodies to document impacts of massive data aggregation and algorithmic outputs, aiming for social and economic justice	Develop global and intersectional analyses of how datafication and corporate monopolisation impact justice and democracy and articulate the common good
	Develop, facilitate, and support participatory action research in collaboration with and within local communities	Build tech literacy in local communities and their capacity to use data on their own behalf, or to resist its use against their interests
	Undertake research to expose harms of data-intensive tools and infrastructure, and identify strategies for democratic data governance	Add to knowledge about the political economy of data. Enhance public awareness of harms of poorly constrained datafication
	Fund individuals, organisations, and institutions engaged in the activities described in this table	Redistribute power over data, and power in decision-making about data, away from Big Tech companies towards the political community
	Engage in litigation to enforce existing data governance frameworks	Engage in litigation to enforce novel data governance frameworks

SOURCE: E. Tucker and this book’s other authors

Research in this area often uses ethnographic methods to reveal how people invent new practices that – even if temporarily – enable them to transgress dominant algorithmic systems (Bonini & Treré, 2024). Some research casts doubt on the efficacy of certain resistance practices, arguing, for example, that “it is questionable whether an average user can actually accomplish the task of understanding and recognizing all risks and challenges related to privacy in an increasingly complex and ever-changing media environment” (Masur et al., 2021: 10). In the following, we provide examples of ways of defending individual rights, including epistemic and collective interests (Horowitz et al., 2024).

Individual digital self-defence strategies

A basic practice of algorithmic resistance available to individuals is digital self-defence. This is increasingly taught globally and practised by activists as part of the work of political organising. Community-based organisations are spreading digital self-defence resources to give people (including young people and children) the means to protect themselves from some of the most acute consequences of data surveillance. Digital self-defence practices may include managing privacy settings, improving the security of passwords, accessing

the Internet through a virtual private network (VPN), using messaging apps employing end-to-end encryption, and removing one's own data from public websites and – where possible – from the custody of data brokers. Defences may include changing the ways platforms are used or opting out of purchasing certain digital products. Some digital self-defence curricula encourage opting out of most social media, or they suggest more secure platforms for online engagement, and basic data literacy training is often a component of digital self-defence training. While digital self-defence focuses on the personal security of individuals, and not on systemic change-making, creating the possibility for improved digital security is an important precondition for larger-scale political engagement, which requires resistance to the marketing claims and promises of Big Tech companies (SSD EFF, 2023).

Public interest alternative news media

Resistance to the harms of mis- and disinformation by public media – including independent public service media – may take the form of building skills and developing ethical data practices or initiatives to counter mob efforts to censor speech online when it threatens to silence journalists and puts their lives at risk (Nechushtai, 2023; Waisbord, 2023). Alternative ways of providing news are not always aligned with equity and inclusion, but alternative news media can operate as a counter to mainstream media when these organisations respect human rights and democratic values. They may engage with local, national, and international human rights bodies to expose and resist the exploitative data practices of Big Tech companies, and they are often well-positioned to investigate discriminatory data practices. The content they produce can help to build public awareness of poorly constrained datafication practices and mobilise efforts to combat the harms of mis- and disinformation (Reiter & Matthes, 2023).

In Latin America, the rise of alternative news media outlets is attributed partly to the necessity for democratic political communication to address information and power imbalances. Alternative news media have historically played a role in countering transnational communication patterns and cultural imperialism, often disseminating counter-information and dissenting against the establishment to facilitate political change. These news organisations are operated by professional journalists who follow the same professional standards and practices as mainstream media (Harlow, 2022). Alternative news media and the use of social media for political participation varies across Latin America (Mitchelstein et al., 2020). For example, in some areas, the main channels for the dissemination of mis- and disinformation have been “alternative” social networks (e.g., X/Twitter, Facebook, and Instagram) and closed social networks and messaging services (e.g., WhatsApp, Telegram, and Facebook Groups). These closed networks have been found to account for a large share of mis- and disinformation, with legacy news media contributing a much lower share (Guallar et al., 2022). Alternative ways of providing news

are not always consistent with equity and inclusion and can be hyperpartisan and disseminate mis- and disinformation (Recuero et al., 2022).

In the United States, with charges of bias against legacy news media by the Trump Administration and an Executive Order ending support for National Public Radio (NPR) and the Public Broadcasting Service (PBS), which previously had very small amounts of support (The White House, 2025b), calls by researchers for a fundamental rethinking of the legal foundations of the news media in the country to support press freedom by respecting the obligation of the government to provide for access to high-quality media have become more urgent (Pickard, 2024). Strategies to enable the news media to contribute to informed public discourse do need to draw a distinction between alt-right sites (intensely engaged in spreading mis- or disinformation), organisations practising journalism with a partisan bias, and those engaged in journalism in support of the public interest and democratic debate (Bennett & Livingston, 2018).

When battles over “truth” intensify, mainstream state-owned media can be branded as “fake news” outlets, while private and independent (alternative) news media may be seen as bearers of “truth”. In sub-Saharan Africa, for example,

media repression through the enactment of draconian pieces of legislation and the brazen capture of legacy media infrastructures by political and economic elites have been followed by the mushrooming of fake online news sites, faceless social media influencers, pseudonymous social media accounts, and coordinated circulation of false and misleading news information through mostly Twitter, Facebook and WhatsApp. (Mare et al., 2019: 5–6)

Another example is India, where alternative news media are positioned in opposition to legacy news media, which is striking because the former support the right-wing politics of the ruling Hindu nationalist Bharatiya Janata Party (BJP) instead of the opposition (Chadha & Bhat, 2022). The BJP-led government has also embraced the use of alternative social media platforms in response to the removal of accounts of prominent right-wing leaders from mainstream social media for violating platform policies (Bhat, 2021). Whether alternative news media initiatives aimed at increasing the integrity of information contribute a vibrant information ecosystem depends on the context and generalisations should not be drawn without evidence.

Community collaborative strategies

Community collaborations provide a basis for resistance to exploitative data practices by Big Tech companies when they produce or collect data that is responsive to community needs. Initiatives in this area include a wide range of communities, from the very local to the municipality, and they take a citizen-first rather than a technology-first approach to data governance.

Indigenous communities, for example, have been establishing alternative data practices through the Data Sovereignty Movement (Our Data Indigenous, n.d.). To maintain control and autonomy over their own data, they have built their own physical infrastructures for telephone and Internet services, and developed apps, browsers, streaming platforms, and messaging services that operate in their own languages (Carroll et al., 2019; Dibenedetto, 2021). Māori activists in New Zealand have used audio data from recordings of people in te reo Māori, the Māori language, to preserve knowledge of te reo Māori for future generations. Community members have rejected offers from tech companies to incorporate te reo Māori audio data into mainstream apps such as DuoLingo and Google Translate, arguing that “our data would be used by the very people that beat that language out of our mouths to sell it back to us as a service [...]. It’s just like taking our land and selling it back to us” (Hao, 2022: para. 29, quoting Peter-Lucas Jones, a project leader). Indigenous communities have also drawn attention to how the predictions of algorithms, such as Google Search’s Autocomplete, treats gender and political keywords in languages such as Amharic, Kiswahili, and Somali in ways that amplify power imbalances (Chonka et al., 2023).

Municipal-level community strategies may go beyond regional- or national-level data governance arrangements, often emerging in direct opposition to the imposition of so-called smart city initiatives that build digital surveillance technologies into the public landscape to collect data for purposes from policing to development planning (Galič, 2022; Rosol & Blue, 2022; Purandare & Parkar, 2020). Smart city initiatives are often driven by multinational companies that stand to profit from embedding their intellectual property in municipal infrastructures (Cooke, 2020). The global Cities Coalition for Digital Rights (n.d.: para. 1) was formed in 2018 to promote and defend “digital rights in urban context through city action, to resolve common digital challenges and work towards legal, ethical and operational frameworks to advance human rights in digital environments”. Many of its 60 member cities have introduced policies and practices designed to shape the way that data is produced and used and to involve citizens directly in decisions about what digital life will look like.



Barcelona is a city that successfully pivoted from a corporate-centric to citizen-centric smart city model (Fernandez-Monge et al., 2024). In 2015, after the election of housing and human rights activist Ada Colau as Mayor, the city embarked on a process of re-envisaging policymaking about data and digital technology as participatory democracy. Barcelona City Council developed a Barcelona Digital Plan, enabling the city to introduce policies and practices for data and the digital infrastructure

(Bria, 2018). These included opening up the digital architecture by developing open standards and prioritising open-source technology over proprietary systems; integrating local providers into procurement; treating data as a shared resource (owned and controlled by citizens rather than treated as a commercial asset) by preserving citizen data ownership in city contracts with digital providers; and using technology to facilitate civic participation in municipal policymaking. In 2021, the city introduced its AI Municipal Strategy with four key principles: AI may be used in the generation of automated recommendations, but may not be integrated into decision-making systems; algorithmic models and digital databases should be transparent and auditable; robust accountability and liability regimes apply when the use of AI tools results in harm or loss; and strict procurement clauses protect municipal control of any private or externally provided AI product (Ajuntament de Barcelona, 2023).



Municipal policies to improve privacy, limit surveillance, or place guardrails around the acquisition and use of GenAI data are increasingly common in the United States. Seattle was one of the first cities to enact a Surveillance Ordinance in 2013, which required city departments to submit guidelines for how they planned to use digital technologies, and what types of data would be produced (Stevenson, 2016). Four years later, it was amended to remedy the absence of language addressing analytic software. The revised ordinance included provisions for holding community meetings prior to city council approval of surveillance technology (hardware and software) acquisitions for use by city departments. New York City's Public Oversight of Surveillance Technology (or POST) Act in 2020 focused more narrowly on the police department, requiring it to produce impact reports and use policies for surveillance technologies that include algorithmic tools (NYPD, 2024). The POST Act was criticised as ineffective by groups that advocated for its adoption, largely because it entrenched a bureaucracy within the police department that legitimises surveillance technology without imposing any real mechanisms for independent enforcement (Dyson, 2023). A growing movement aims to limit the use of specific surveillance technologies at the municipal level, and 21 municipalities now have ordinances banning or restricting the use of facial recognition technology by certain actors or in designated contexts (McConvey, 2024; Tate-Mosely, 2023).

Social entrepreneurs and community-controlled technologies
Digital resistance also involves technologies that are owned and controlled by local communities, usually for purposes specific to the needs of those

communities. Social entrepreneurs may partner with local communities to develop applications based on non-extractive data practices that challenge the dominance of corporate data governance models and offer alternatives for individuals and communities. An example is Alt (Alternativa Laboral Trans), a worker cooperative in Argentina owned by trans and non-binary people which offers digital design and development services using a non-data extractive business model that relies on open-source software. This allows clients to maintain as much control as possible over the digital afterlives of their work and provides support for digital tools for the protection and education of its community members (Mejias & Couldry, 2024).

The National Digital Inclusion Alliance (NDIA) works in the United States with local community organisations to promote digital equity, inclusion, and literacy, especially in areas underserved by corporate broadband providers. Initiatives aim to empower local community actors to take an active role in pursuing digital equity. Through programmes like the National Digital Navigator Corps, it provides personalised tech support and digital literacy training, helping to bridge the digital divide (Menon, 2024; NDIA, n.d.). In inner-city Milwaukee, Wisconsin, a partnership with university researchers was reached to create their own geographic information system (GIS) to enable community members to engage politically based upon their access to relevant information. The aim was to create a database that would help to identify property sales in their neighbourhood, absentee landlords who had abandoned properties that were deteriorating, and enable delinquency and building code violations to be taxed (Ghose, 2001). Relying on a combination of public datasets, local data, and the knowledge of local community members, a database with a neighbourhood map interface was built, allowing users to retrieve data necessary to participate in neighbourhood planning processes (Ghose, 2001).

Social entrepreneurs and nongovernmental organisations also work with local communities to use AI tools or build digital networks that serve communities, which are controlled so as to avoid participating in data extractivism. The Distributed Artificial Intelligence Research (DAIR) Institute was established to undertake research that benefits “communities which are typically not served by AI and create pathways to refuse, interrogate, and reshape AI systems together” (DAIR Institute, n.d.). It works to articulate alternative principles for data use and enact these in their own digital tools or resources. Examples include a tool using computer vision and satellite imagery to visualise the impacts of spatial apartheid in South Africa (Tsanni, 2024); the use of machine learning to analyse the history of racial justice protests in the United States (Oliver et al., 2022); and a wage theft calculator that estimates how many wages are lost (stolen) as a result of a surveillance technology called Mentor (Williams, 2023).



An example of building community-controlled networks is the work of Uganda Flying Labs (part of a larger humanitarian organisation with projects in 32 countries). Drone technology is used for disaster relief, services planning for refugees, and agricultural monitoring and community development. Its data was used to identify evacuation points during the 2019 Bushika landslides in eastern Uganda, to coordinate Covid-19 responses in local refugee settlements, and to help coffee farmers respond to extreme and unpredictable weather patterns. Flying Labs coordinates with local government to implement projects, but the local government does not own the data. The use of drones for development-oriented initiatives does require private sector investment, but “the data are owned by the people who commission the research rather than the Flying Labs or the broader community” (Horst et al., 2024: 1372). This, along with the lack of resources available to establish servers for storing and managing data, means the data that Flying Labs collects cannot be co-opted for other purposes.



In Cape Town, South Africa, a nonprofit organisation – Violence Prevention Through Urban Upgrading (VPUU) – developed its own data infrastructure to freely provide a licensed community wireless network serving over 65,000 people. V-NET is a mesh network with local Internet access to support community development projects including educational and social programmes, public works projects to improve local infrastructure, and advocacy for better service delivery. Apps are used to collect data, and community members are trained to track the maintenance of public facilities, analyse community needs, and use the data as evidence to advocate with local government (Blake et al., 2023; VPUU, 2019).

In the Middle East and North Africa, the nonprofit organisation Majal (2024a), founded in 2006 by Bahraini social entrepreneur Esra’a Al Shafei, develops digital spaces to foster “safe and accessible expression, association and communication” for underserved and underrepresented communities. One of its platforms, CrowdVoice.org, collects, curates, and contextualises crowdsourced data about global protest and social justice movements and serves as a resource for activists and journalists. Majal is co-founder of the Numun Fund, which aims to “seed, resource and sustain feminist and women/trans-led groups who engage with technology in their activism” (Majal, 2024b; see also Skalli, 2023).

National-level decentralised data governance frameworks

Even when many governance institutions are captured by industry interests or communities become dependent on Big Tech company infrastructures and services, communities can engage in organised resistance by dissenting from corporate- and politically-controlled datafication. Alternative data governance models offer promising departures from centralised data governance systems by promoting local ownership, control, and benefit-sharing. India's Non-Personal Data Governance Framework and digital citizenship initiatives in Latin America, for example, provide a solid foundation for these models, which foster a sense of data ownership and empowerment and help build trust and cooperation by ensuring that communities are the primary beneficiaries of the data they generate, potentially leading to more equitable and sustainable development outcomes. India's framework focuses on nonpersonal data generated in rural and urban areas. Data generated by rural gram panchayats (village councils) and urban municipalities is owned by local bodies, referred to as data stewards, with the aim of democratising the benefits of data exploitation and giving local communities a say in how their data is used. Data stewards or trustees serve as data guardians, managing and regulating access to this data to ensure data privacy, secure data rights, and foster a transparent decision-making environment (Bailey et al., 2020; Jindal & Nigam, 2020). These data governance frameworks support a variety of community-based products and services (Kumar et al., 2023; Lanier & Weyl, 2018; Micheli et al., 2020; Singh, 2020; Verdegem, 2021; Olorunju & Adams, 2024).

Approaches to data governance based on “commons” principles offer an alternative to proprietary models, often by leveraging the affordances of cloud-based software platforms using open governance that allows data to be managed, analysed, accessed, and shared within a community. Rooted in principles of Open Access and collective benefit, and drawing on **practices** in free and open-source software (FOSS) communities, collective data management strategies can support communities in political claims-making. Data is stored, actively managed, and analysed using shared tools, reducing redundancy and enhancing the quality of data analysis. These initiatives are a promising alternative to centralised models of data management and AI system development, although there is always a risk that they become conflictual if they must intersect with commercial entities and states that seek to appropriate their data resources (Berdou, 2011; Birkinbine, 2018; Dalle & David, 2007; Powell, 2015; Zygmuntowski et al., 2021).

Other examples include community data hubs, which operate as localised data centres where data is stored, processed, and managed, with oversight from local data stewards, and serve as centres for innovation and learning,

with data literacy and data rights training programmes. In some cases, data infrastructures are using blockchain or other decentralised technologies such as IPFS (InterPlanetary File System), decentralised identity systems, and smart contracts to ensure that data transactions are secure, transparent, and accountable to local communities. Local data analytics services are being developed to analyse local data and provide insights of benefit to communities (e.g., agricultural advisories in rural areas using local climate and soil data, or urban planning tools using mobility data to improve public transportation systems). Data lockers and intermediary organisations can enable people to control their own data, although time and knowledge are needed to make judgments about when to release data and to whom. In initiatives involving data cooperatives, women, farmers, and others are being enabled to retain control of their data and the rewards that flow from its use, with public data infrastructures being developed in sectors such as commerce, finance, health, education, and agriculture.

These strategies are facilitating data sharing within and between communities and supporting collaborative projects (Menon, 2024). Many offer a way to dismantle the concentrated power of Big Tech by using the computing infrastructure to create a communal utility (Verdegem, 2022). Strategies involving AI systems and data depend on access to growing amounts of computing power, and there is a range of national initiatives to deliver public access to “compute” power – both publicly and privately owned. Many of these strategies aim to support scientific research and smaller and medium-sized companies, but there are also warnings against capture by Big Tech and the need to support a variety of initiatives so that many organisations can benefit (Davies & Vipra, 2025). These alternative data governance approaches may be resource-intensive with limited scalability and require periodic updates and hardware and software maintenance, which can be costly (Jackson et al., 2025), and they may lag in adopting the latest global technical advances. In the Global Majority World and elsewhere, however, these strategies aim to stop large technology companies from being the “privileged providers of social solutions” (Magalhães & Couldry, 2021: 354; see also Adel et al., 2023; Gurumurthy & Chami, 2022; Nyalety et al., 2019; Omar, 2023; Page et al., 2023; Vayadande et al., 2024).

Civil society organisations, researchers, and philanthropic organisations

The examples in the preceding discussion underscore the importance of local control and community-specific solutions in initiatives to address digital inequalities and injustices. Unlike governance that merely seeks to balance corporate investment priorities with collective interest, these initiatives emphasise local capacity and agency. Communities using tailored digital solutions are addressing their unique challenges and seeking to ensure that technological advancements do not perpetuate inequities. Most documented

examples are small scale and local, but they demonstrate the possibility of, and provide inspiration for, meaningful collective action in contexts where it is difficult or impossible for most people, let alone the vulnerable, to access formal political processes.

Civil society organisations, including unions, are engaging in power struggles to establish systems of social support for communities harmed by commercial datafication (Blum & Uldam, 2025; Graham & Dittus, 2022; Graham & Ferrari, 2022). In many cases, they are collaborating with researchers to mitigate harms and remove barriers to collective organising for those who are most negatively impacted by datafication. They are helping to build local community capacity to use data on their own terms and to resist data uses that are not in their best interests. Their work exposes harms of data-intensive tools and infrastructure and identifies strategies for data governance that are democratic in practice, not just in stated policy. They produce knowledge about the political economy of data, enhancing public awareness of the epistemic and material harms of poorly constrained datafication.

This work is critical of technology-centred approaches, and its focus is on data justice. It helps to amplify the reach of social movements that are supporting democracy and capacities to control the uses of data and technology, including AI systems, in line with human rights commitments and by reducing inequalities. With funding from philanthropic organisations, crowdsourcing, or drawing on (scarce) university funding, they contribute to local (sometimes global) initiatives for establishing ways to redistribute power over data – and over decision-making about data – away from Big Tech companies (see Cammaerts, 2018, 2024; Crawford et al., 2014; Dencik et al., 2016; Dencik & Leistert, 2015; Hepp et al., 2022; Milan & van der Velden, 2016; Ó Siochrú et al., 2024; Timcke & Hlomani, 2024).

A variety of forms of participatory action research by these organisations in collaboration with, and within, local communities is helping the resistance against corporate datafication by identifying how concentrations of power in the hands of large companies and states is resulting in intrusive datafication and surveillance, sometimes involving biometrics (Munoriyarwa & Mare, 2022), facial recognition technologies in Brazil (Ramiro & Cruz, 2023), inappropriate uses of technologies exported from the Global North to African countries (Sheombar & Skelton, 2023), AI-enabled surveillance and data collection used to monitor migrants and refugees (Napolitano, 2023), or “smart spying” in the United States (Moran et al., 2023). As a “new aesthetics of surveillance” takes shape (Beiguelman, 2021; Calderaro & Blumfelde, 2022), this work focuses on “the absences, the silences and the forgotten and ignored people and regions of the world” (Gillwald & Wavre, 2024: 34; RIA, 2023a). The resistance strategies discussed here do not assume that the “good” is always embodied in technology or data

(UN, 2020), as is the case with some alliances of governments and the private sector to deliver “digital public goods” for the Global Majority World (DPGA, 2023). In contrast to initiatives that do not empower local communities to control and own their data, good outcomes are more likely when initiatives are not driven by technology, but instead by explicit community choices about the design and operation of technology.

Conclusion

This chapter began with a review of research on the public’s awareness of the risks and harms linked to mis- and disinformation and of how online content is moderated, and whether policymakers understand the impacts, especially in the context of elections. People have varying degrees of confidence in their capacities to identify AI-generated information, and support for interventions by Big Tech companies and governments to tackle mis- and disinformation varies by country context. There is no “silver bullet” response that is likely to increase public trust in the information that circulates online. It is clear, however, that multistakeholder dialogue is an essential requirement if inclusive responses to mis- and disinformation are to be put in place.

We reviewed a variety of measures to combat mis- and disinformation, indicating that a single approach is unlikely to be feasible or desirable. This applies to efforts to inform the public as well as government actions through legislation and the introduction of codes of practice that the Big Tech companies are expected to adhere to. Evidence indicates that certain content governance methods, such as fact-checking, have limited effects on people’s responses to mis- and disinformation, and that these approaches must take account of changes in techniques and practices over time. In addition, although the use of AI tools to support automated fact-checking and content moderation may be inevitable, recognising that no algorithm is neutral and the need for “meaningful” platform transparency are crucial if the Big Tech companies are not to operate as the “new governors” and if risks of censorship are to be reduced. This discussion highlighted the need for independent comparative research and work to ensure that data is available to support research. Overall, the analysis confirmed that there is little clarity about the most effective intervention points where content moderation can have positive impacts that help to encourage political engagement and protect human rights. There is always a risk that policy enforcement to combat mis- and disinformation will lead to the suppression or violation of human rights.

This chapter also examined strategies that aim to resist corporate datafication, focusing on how both individual and collective measures help to reimagine new pathways to achieving local or decentralised control over the governance of data and digital technologies. The main focus of research in this context is to reveal how knowledge, structures, and practices can be

combined with the affordances of technology in ways that shape what counts as “good” knowledge and practice. Individual and collective dependencies and inequities in information ecosystems manifest in the form of discriminatory decisions in the treatment of people by, for example, gender, race, ethnicity, disability, class (caste), and language community. The focus here was on practical action and the agency of a variety of actors to achieve data justice by devising beneficial uses and frameworks for governing these technologies. Attention was given especially to the need to tackle the epistemic privileging, which downplays the agency of diverse actors to build information ecosystems that are aligned with their needs and cultures. We also emphasised the need for a decolonisation of research approaches and traditions to avoid exclusions from, and inequitable inclusions within, information ecosystems.

Conclusion: Information ecosystems and troubled democracy

“Information and freedom are indivisible”
(Kofi Annan, 1997: para. 20)

Introduction

Meta’s Mark Zuckerberg proclaimed in January 2025 that less intervention into what people find on its social media platforms is the new “best practice” (Kaplan, 2025). The company had decided that its guidelines to identify harmful mis- and disinformation were too complex, and its algorithms to detect harmful content too prone to error. Announcing that they were abandoning independent content fact-checking (in the United States initially) and that their community content guidelines would be recast in the name of promoting debate on contentious issues like immigration or gender identity, Zuckerberg acknowledged that “we’re going to catch less bad stuff” (Chow, 2025: para. 4). United States–owned social media companies that are subject to European Union platform regulations have claimed to be unfairly targeted, characterising the Digital Services Act, for example, as a case of “institutionalising censorship” (Blenkinsop, 2025: para. 3). President Trump has said that if the enforcement of European Union legislation results in fines levied on Big Tech companies, this amounts to “overseas extortion” (Moen, 2025). In Africa, shutdowns of the Internet (reported by Access Now and other civil society organisations) ordered by governments, militia, and other non-state actors numbered 21 in 15 African countries in 2024, the greatest number ever recorded, while the total number of shutdowns across 54 countries globally was 296 (Egbejule, 2025). These were mainly responses to conflict, protest, or other forms of political instability. Newsguard, which provides information reliability ratings, has identified 1,271 unreliable AI-generated news and information websites that operate in 16 languages (Sadeghi et al., 2025). These developments are only a small part of the context in which information ecosystems are developing and in which these developments are troubling democracy.

The research we reviewed in the preceding chapters helps to illuminate why it is challenging to creating inclusive, safe, and secure information ecosystems that foster the consumption of reliable information, and which protect and promote human rights and the collective or common good. Research findings

cannot predict the future of information ecosystems with certainty, but they do reveal consensus and dissensus on the values and decisions that are determining how these ecosystems are developing in their different contexts, and on the values and norms that should guide their future. We have focused on research about what is known about the changes in the news media, AI systems, and data governance components of information ecosystems, emphasising how these ecosystems incentivise mis- and disinformation and how dependence on Big Tech companies and their exploitative datafication practices is produced. Our analysis confirms that AI systems are being deployed in ways that infringe on individual rights and on collective interests in fairness and justice, even as they bring benefits for some individuals and groups.

Throughout, we have sought to emphasise that the integrity of information within information ecosystems is the result of human decisions and actions and that it is informed by historical and contemporary inequalities in distributions of power. It is these inequalities that condition whether the development of information ecosystems has the potential to contribute positively to public debate and democracy. In our analysis we have treated growth in the production and circulation of mis- and disinformation as symptoms of complex changes in society and as amplifiers of these changes. We have evaluated the research in our corpus through this lens to demonstrate what is known about the determinants of this reciprocal relationship and about the effectiveness of diverse governance approaches. In the following, we highlight what we can learn from this large-scale and critical assessment of research across multiple disciplines.

Epistemic difference and inequities

Our analysis confirms a neglect of the distinctive characteristics of information ecosystems, especially in the Global Majority World. It demonstrates a very strong bias towards research based on the experience of platformisation and datafication in the Global North, and especially, the United States. This strong Eurocentric and Western bias towards research conducted by scholars in the Global North, and about how information ecosystems are shaped by institutions within the Global North, means there is a lack of attention to how diverse social, cultural, political, and economic values, even if they are contested, shape the risks and harms associated with information ecosystems. This bias is particularly visible in studies of the effects of mis- and disinformation on individuals, and in research on content governance and platform governance. A review of studies of digital platforms also found this bias, showing that countries with the largest proportion of papers on the topic were the United States, China, and the United Kingdom, and, as a region, the European Union (Ha et al. 2023).

Epistemic injustice is present when only certain kinds of knowledge are privileged. Tackling bias and injustice requires critical research and a much greater effort to decolonise theory and empirical practice. This is essential

if knowledge about and the experience of data-intensive economies in the Global Majority World is to inform strategies for governing information ecosystems. Our analysis confirms the need for creative approaches to ensure that Eurocentric or Western knowledge is not the unquestioned guide to future developments of information ecosystems. This departure from the hegemonic knowledge of the Global North is essential so that the Global Majority World is not treated as a passive recipient of knowledge and digital technologies, and to address the risk that information ecosystems exacerbate existing inequities and injustices. Many studies of the risks and harms associated with information ecosystems fail to take the absence of, or uneven distribution of, meaningful Internet connectivity (affordability, skills, and outcomes) into account. Findings about developments in the Global North are often implicitly assumed to be generalisable beyond the Global North. Larger-scale comparative survey studies do extend beyond the Global North, but they examine impacts mainly at the country level.

If the *Global Digital Compact's* goal of promoting diverse information ecosystems that embrace international human rights commitments is to be achieved, the structural and epistemic biases of research on information ecosystems are in urgent need of attention. Questions include whether investment in digital infrastructures and services is entrenching problems experienced in higher-income countries with their advanced digital infrastructures and whether diverse data sources are available to inform research. Comparative research on state and co-regulatory governance institutions and processes within the Global Majority World are also needed, as is comparative research on bottom-up governance initiatives involving civil society organisations and local groups. Many of these initiatives are present in the Global North and in the Global Majority World, but too little is known about them.

Conceptual framings, research designs, and methodologies

As we indicated above, research evidence confirms that the affordances of information ecosystems (including AI systems) are often implicated in failures to protect human rights. Our review illustrates that, beyond high-level statements of principles, the problem of how to tackle risks and harms effectively is framed in diverse ways. When the components of information ecosystems are understood as being interdependent, however, research is more likely to reveal the complex factors that influence the production and circulation of mis- and disinformation and the way that unequal power operates through digital infrastructures and services, and through political institutions. Even if there is broad agreement at a high level that fostering information integrity is important, we found disagreement about what this concept means, what should be assumed about normative goals, and how these are, or should be, translated into practice at local, country, and regional levels. Our analysis has pointed to a striking use of multiple definitions of concepts, with illegal or harmful

information defined differently across jurisdictions, and we found numerous and vague uses of the term “AI”. Operationalisations of these concepts differ across both research traditions and regions and countries, and in many cases, these differences are not explicitly noted.

It is also relatively rare for several of the components of information ecosystems to be studied at the same time, and there are disagreements about whether the object of interest should be the ecosystem or the public sphere. In addition, we found few signs of learning from earlier mainstream and critical studies of the “information society” or the “knowledge society”, both of which referred to contexts with which researchers and policymakers are grappling today when the information ecosystem concept is used. In some studies, the concept of information integrity is seen as being too open to interpretation of what is “good” information, and in others, it is the “public worthiness” of information that is of interest, with a focus on informed public discourse. Overall, we found considerable slippage in the entries in our corpus between concepts of “data”, “information”, and “knowledge”. This matters when the problem is to investigate not only the materiality of technologies and their impacts, but also the epistemic issues around whose perspectives matter and shape the development of information ecosystems.

Conceptual variations are common in all academic research. However, our review indicates the need to build a more holistic understanding of the determinants of changes in information ecosystems. This requires an effort to build bridges between the humanities, social sciences, and computer science and an investment in large-scale, comparative studies that focus on the factors that shape the whole lifecycle of the production and circulation of mis- or disinformation, the moderation of this information, and its consequences for democracy.

Research design and methods

Research on information ecosystems often starts with an interest in information (or news media) and its impact on individuals. This contrasts with research that begins with questions about the determinants of inequalities and how monopolistic Big Tech companies, undemocratic institutions, and failures to respect international human rights contribute to democratic instability and to the role of mis- and disinformation within this context. In both research traditions, there may be a concern with risks and harms, discrimination, or damage to the public sphere, but in the first tradition, predictive studies are typically designed to identify links between mis- and disinformation, attitudes and behaviours, and political polarisation. In our corpus, we found that many of these studies draw upon theories of media effects that have been questioned by numerous scholars and are insensitive to institutional and market conditions and distorted power relations. When these conditions are the focus of research, it is more likely that the causes of social discord and distrust, including unequal power and conditions that enable Big Tech

platforms to flourish, will be found to provide a more robust explanation of the impacts of mis- and disinformation.

Our assessment of research demonstrates the need for a much greater emphasis on the socioeconomic, political, and cultural conditions that become institutionalised within information ecosystems which then foster mis- and disinformation and democratic instability. Multidimensional research designs are needed that focus on relationships between the components of information ecosystems attending both to the effects of mis- and disinformation on individuals and to the context that incentivises production and consumption of this information. Longitudinal, comparative studies are needed with global coverage focusing on the companies – smaller and larger – involved in datafication, on what publics believe should be done about illegal and harmful mis- and disinformation, and on which actors are represented in deliberations about how to govern information ecosystems.

There are many studies on the effectiveness of initiatives to counter mis- and disinformation employing quantitative methods and conducted in experimental or quasi-experimental settings. However, studies of the use of AI tools are prone to neglecting the complex factors that may explain the effectiveness of these measures. This can be revealed when qualitative methods including interviews, focus groups, or storytelling, and thematic-, discourse-, content-, or document-analysis are used to elicit deeper insight. Greater use of qualitative methods could elicit insight into power disparities – explicit and hidden – that influence choices about the design and deployment of AI systems and the agency of children, adults, and communities in deciding how they engage with digital content. For example, qualitative research on AI-driven mis- and disinformation campaigns can provide fine-grained insight into how these campaigns operate, what drives government actors and other influencers, why information is shared, and why trust in news media varies. Greater use of interdisciplinary approaches and mixed methods can yield better insight into both the affordances of technology and the practices of states, companies, and other actors.

Research also needs to encompass a larger number of digital platforms, including those originating outside the United States and China. Measuring the scale of mis- and disinformation and the severity of its impacts will remain challenging without access to data that can reveal how business models and technology combine to create incentives for the production and sharing of mis- and disinformation. A wider range of researchers, including civil society organisations and those in the Global Majority World, needs access to useable data if Big Tech companies are to be held accountable by those who seek to govern them. Data governance frameworks to enable the provision of granular platform data that can support independent assessments of Big Tech business models and practices are therefore essential. This requires public data infrastructures and resources, especially in the Global Majority

World. When data frameworks to allow access are established, it is essential to monitor whether they are respected and whether researcher independence is safeguarded from political interference and corporate pressure to interpret research in ways that favour the interests of companies or governments. One approach may be to establish a network of multinational and interdisciplinary research centres dedicated to the study of the role of information ecosystems and the implications of mis- and disinformation for democracy.

Future research priorities

In this section, we provide greater detail on the priorities for future research that emerged from our research assessment. We start with the need for research that focuses on how Big Tech datafication business models that thrive on commercial data monetisation are implicated in threats to the integrity of information and to democracy. Establishing improved Big Tech and government accountability as well as transparency is a high priority. We then focus on the need for research to assess the roles and effectiveness of approaches to governing the infrastructure layer of information ecosystems, regulating news media, governing AI systems, and approaches to combating mis- and disinformation. Lastly, the need for research in support of efforts to resist exploitative datafication practices is discussed, focusing on work that can assist in devising alternative data governance frameworks to the centralised ones that are being put in place and in facilitating inclusive democratic dialogue aimed at achieving data justice. Across all these areas, it is essential that research is undertaken in a way that is inclusive and able to capture the experience of the Global Majority World.

Datafication-for-profit business models

Research shows how powerful monopolistic Big Tech companies are able to determine what data is produced and how it is produced. Research is limited, however, by restrictions on access to data and, increasingly, by further political pressures on researchers. This is especially so when research reveals how business models and the features of AI systems result in discrimination by, for example, gender, race, ethnicity, disability, class (caste), or language community – not only for individuals, but for whole population groups. The digital platform business models, which drive data-intensive economies and labour markets, are creating incentives for the production and circulation of mis- and disinformation. Even if research is unable to identify the exact causal relationships between the affordances of technology and these outcomes, it can reveal enough about how data extraction practices operate to provide the evidence needed to hold Big Tech companies and governments to account. This work can also help to build awareness and to mobilise initiatives to support participation in democratic change processes.

All sectors are becoming dependent upon Big Tech's data-intensive products and infrastructures (e.g., from healthcare to education, transportation, the news media, etc.), and research is needed to reveal how this is progressing and its consequences. This means monitoring Big Tech self-regulatory practices to expose when they abandon, or arbitrarily change, their content governance processes or when they weaken their privacy and data-protection policies. This kind of evidence is essential to inform efforts to counter harmful exercises of corporate power. It includes work to track third-party data-sharing agreements, the incidence of data breaches, and efforts to address known algorithmic biases. It also requires work on how personalisation algorithms are used to make platforms attractive targets for mis- and disinformation campaigns and to evidence Big Tech complicity in intervening in political campaigns.

Governing information ecosystems

There is considerable research on legislative approaches to governing information ecosystems. However, not only is coverage of countries beyond the Global North uneven, but there is also a gap between governance goals and norms and their implementation; this is especially so in efforts to mitigate the harms of mis- and disinformation. Research is needed to identify the contexts in which governance interventions are most likely to be effective and which combinations of approaches may help to improve the integrity of information. We have highlighted research that evidences how the regulation of both legacy and online news media in some contexts is resulting in censorship or political leverage over news media organisations, and we have drawn attention to work that shows why privacy and data-protection legislation are not panaceas for mitigating the harms of exploitative data practices.

Research is needed to assess whether governance mechanisms are aligned with international human rights (including the rights of the child) as well as with collective interests, to examine industry compliance, and to highlight examples of contexts where policies favour the interests of Big Tech companies over those of other stakeholders. Specifically, what rules need to be in place to assure that Big Tech operations are “meaningfully transparent” and to monitor how they are being enforced? Are Big Tech reports sufficiently transparent, especially regarding content moderation and curation (including personalisation algorithms), and are inclusive mechanisms for consulting individual users and collective organisations in place? A weakness of existing research is that, while it does try to examine the impacts of various governance approaches on individuals (with often contradictory results), it grapples less often with whether these approaches are effective in helping people safely navigate information ecosystems or in enhancing their resilience to mis- and disinformation.

The infrastructure layer

This research assessment was not intended to examine research on the contribution of developments on the infrastructure layer in depth, but it did identify the need for research on how network neutrality policies and zero-rating contracts are being used to limit access to information in ways that can curtail the diversity of information available to people in many areas of the world. This research is crucial because the impacts of these policies and data-pricing practices condition whether informed participation in the public sphere is feasible. These policies affect who is connected to the Internet and whether governments can disconnect people or limit information flows, especially during elections or political unrest. The studies that are available indicate that these policies have markedly different effects in different national (and local) contexts, and research is essential to understand pressures leading to the fragmentation of the Internet and the implications of policies that aim to strengthen national digital sovereignty.

News media and trust

In the case of research on transformations in the news media industry, not only are there numerous definitions of news, but we also found a strong emphasis on online news with a tendency to neglect the role of legacy news media and the contributions of different “genres” of news to the public sphere. There is a strong focus in some of the literature on content produced by professional journalists, with less attention given to other actors who are motivated to produce mis- and disinformation or to the lessons that could be drawn from research on the history of propaganda. Research on changes in the news media industry and in media freedoms must encompass countries in the Global Majority World, taking account of a wide range of Big Tech platforms and actors (state/private and foreign/domestic) and their motivations if we are to understand the changes in trust in the news media, news avoidance, and resilience to mis- and disinformation.

As the news media industry becomes increasingly dependent upon digital platforms provided by Big Tech companies and on personalisation algorithms, this is destabilising news organisations financially and resulting in closures, especially of smaller local news outlets. There is research that focuses on how media ownership rules, regulations, and content governance can suppress content that is critical of state authorities, and the need to protect media freedoms is consistently emphasised. However, the factors that curtail news media independence and explain the different effects of platformisation on legacy and online news media and on the quality of information need to be better understood. This means giving a greater priority to research on changing market structures, the transparency of ad tech markets, agreements reached between the Big Tech companies and news organisations, and means of financing a viable independent news industry in countries around the world.

Our analysis has highlighted how trust in news media and perceptions of the trustworthiness of news media organisations varies across countries, as do the factors that explain news avoidance, selective news exposure, and abilities to discriminate between accurate and false information. It is unclear, however, why trust varies so dramatically across countries and within different groups, which indicates that work is needed on the interdependent factors that explain these results. The research that is needed must go beyond the study of the impact of the news media on individual characteristics to explore what factors contribute to both healthy and unhealthy forms of scepticism towards information sources. It is clear that different types of mis- and disinformation impact the public sphere in a variety of ways, and studies are needed that focus on the actors, in addition to the far right, that are incentivised to weaponise information. The diverse results of studies on the phenomena of filter bubbles and echo chambers and their impact on individuals' political beliefs and attitudes, in some cases leading to political polarisation, also point to the need for interdisciplinary studies and to investigate the circumstances where self-imposed filter bubbles can offer a safe space that helps to protect marginalised groups. In this area, longitudinal studies (with global coverage) using both quantitative *and* qualitative methods are needed.

Studies of how a plural and diverse public sphere and sustained media freedom can be fostered would benefit from more attention to news media formats and to how the use of AI tools by news producers and other actors (governments, bots, or private individuals) is changing news production and circulation. Is the use of AI tools enhancing newsroom efficiency and productivity? Research on transformations in the news media industry also needs to be linked to work on content moderation and efforts to curtail mis- and disinformation, as well as to research on policy aimed at securing the safety of journalists or governing political campaigning and political spending.

AI systems and content governance

Our analysis has highlighted how the term “AI” is used – misleadingly – across the academic literature. It appears as a generic category, which is misleading when the task is to respond to specific risks. Even when research is specific about the object of study, for example, large language models (LLMs), there were few signs in our corpus of work that attends to the social factors that influence system design and implementation. Research on the role of AI systems (specifically machine learning algorithms) and their embeddedness in online content governance systems reveals many human rights violations when these systems are deployed, and there is evidence that automated content governance and algorithmic decisions impact negatively on democratic decision-making when they determine the conditions under which content is seen and with whom it is shared. Technologies, such as GenAI, are challenging individual rights and rights to democratic participation. Much greater insight is needed into how AI systems development and uses are co-evolving with

efforts to safeguard internationally protected human rights and fundamental freedoms. This requires attention to whether these approaches are responsive to specific challenges posed by new actors, instruments, and contexts.

Legal frameworks are essential for defining and removing illegal content, assuring accountability and transparency of problematic content, and establishing rules for algorithmic personalisation systems. AI-enabled algorithms promote or demote content in different ways depending on country context and political influences. These differences across all countries need investigation, with a focus on the societal impact of algorithmic decision-making, including the operation of content governance tools and auditing processes. Specifically, how are AI systems development, the use of automated tools, and pressures on democratic processes related? Answers require research on the whole lifecycle of AI systems development and deployment, and on diverse use and country contexts. What combinations of content interventions are most likely to be effective in mitigating the risks and harms of mis- and disinformation? Research is also needed across countries on the content moderation techniques that are acceptable to online participants and to monitor their effectiveness, including fact-checking, over time and in real-world settings. Researchers call for AI system audits to enhance the transparency of AI systems, but there is less clarity about who should perform audits or how they should be done. Whether algorithmic audits are resulting in “responsible” AI development and use remains an open question, and work is needed to reveal over time how AI systems are impacting disadvantaged and marginalised communities. Research on the “AI divide” (i.e., the disparity between those who can access and effectively leverage AI systems and those who cannot) should also be a high priority, given that the outputs of LLMs cannot be expected to be fully representative or equitably inclusive. AI systems deployed in the news media industry, for instance, will have biases as the result of their design and operation. Mapping this “divide” is needed to identify tensions between the benefits of promoting the use of AI systems in the Global Majority World, where Internet access is absent or very limited, and the risks and costs associated with mitigating risks and harms after AI systems are introduced.

Literacy and informed publics

Our analysis confirms that media and information literacy (MIL) and AI literacy can help to equip people to make sense of the information they encounter in the mediated public sphere. However, conceptualisations of MIL and AI literacy differ across regions. Research findings in this area can be used to help the public understand who is responsible for protecting their rights and to assist policymakers in deciding on responses to the power of Big Tech companies. Which approaches to MIL and AI literacy are effective in the longer term, what skills and training do they involve, and how are they funded in different country contexts? Are interventions enabling children and adults to interpret and critically assess information and to differentiate between legitimate and other

sources of information effective? What progress is being made to standardise MIL and AI literacy frameworks and methodologies to advance comparative evaluations? In the case of AI literacy training, are learners being introduced to knowledge about the fundamental principles of AI systems at all stages of AI systems development and deployment? Holistic assessments of resilience to mis- and disinformation across countries are needed, although literacy training should not be treated as the only solution to risks and harms encountered within information ecosystems.

Governing democratically: Towards data justice

Our analysis has demonstrated that the governance of information ecosystems is not providing sufficiently robust human rights protections. Data justice is not being achieved. We need more systematic evidence on measures that are underway to promote the capacities of communities to question their dependencies on Big Tech company infrastructures and services *outside* government-sponsored policymaking spaces. This is essential, since the problems of online mis- and disinformation and exploitative datafication show few signs of being addressed within the current political and economic order, at least, not in a way that protects collective interests in democracy (Fraser, 2022). In this context, research is needed to underpin radical resistance to dominant technology designs and uses with the aim of promoting justice and equity.

One goal of resistance strategies is to redistribute power away from Big Tech and governments towards local political and social communities. What strategies are individuals and communities developing to reimagine and resist the data extractive features of data-intensive economies? There are several key questions: How can data justice can be achieved by building alternative data governance frameworks? What is the state of knowledge about practices and local solutions, and can these be sustainably scaled up? How effective are initiatives that are underway in devising frameworks for enabling individuals and groups to control their data? Do these initiatives represent the needs of marginalised people and avoid reproducing epistemic injustices when they are introduced?

Research is also needed on how data production and processing operations and services are affecting people's daily lives, and whether AI systems in the Global Majority World are replicating problems with data-intensive economies that are already present in the Global North. What approaches to meaningful deliberation are in use to support novel approaches to data governance, and what is the experience with these data governance frameworks? What are the circumstances where pressure from civil society is successful in treating data governance as a lever for restructuring data markets, protecting against infringements of human rights, and tackling concentrations of power and wealth that jeopardise democracy by adopting democratic deliberative processes that foster international solidarity?

Conclusion

If Alphabet's, Meta's, XAI Holdings', and other Big Tech companies and government ambitions for global competitive success prevail – instead of protecting adults and children from harm and fostering a public sphere where accurate information provides a basis for democratic deliberation – information ecosystems will continue to contribute to the destabilisation of societal order. Our critical research assessment has shown that when the study of today's information ecosystems is framed in a way that encourages the rapid pace of technological change and welcomes the marketisation of new products, with an assumption that risks or harms will be mitigated, it does not reveal how these ecosystems are contributing to democratic instability. In contrast, when research is undertaken that challenges how transformations in information ecosystems are being normalised as a result of decisions by powerful Big Tech companies, and how some states are privileging their own interests in winning a technological race or in perpetuating autocracy, the results can signal the need for changes in institutions and practices to curtail mis- and disinformation, to prevent exploitative data practices, and to uphold human rights.

Research that focuses principally on the impacts of AI-enabled mis- and disinformation on individuals must urgently be complemented by studies of the socioeconomic, political, and cultural conditions that contribute to information ecosystems which destabilise democracy. We have provided a baseline analysis for future learning. We have showcased what is being done and what could be done differently to progress towards equitable and inclusive information ecosystems. We have demonstrated how and why existing research privileges the knowledge produced in and about the Global North. If this bias persists, skewed evidence will inform future policy and practice, contributing not only to democratic instability, but also to a furthering inequality and injustice.

Acknowledgements

This analysis in this book was supported by the International Observatory on Information and Democracy. The book is a revised and updated edition of the Observatory's first cycle report *Information Ecosystems and Troubled Democracy: A Global Synthesis of the State of Knowledge on News Media, AI and Data Governance* published in January 2025. We, the authors of this book, served as scientific director (Robin Mansell), lead rapporteur (Rob Procter), and rapporteurs (Flavia Durach, Matthias C. Kettemann, Théophile Lenoir, Gyan Prakash Tripathi, and Emily Tucker). We express our gratitude to everyone who supported the preparation of the report and especially to Camille Grenier, executive director of the Forum on Information Democracy, and Emma Gruden, project officer, for their encouragement to prepare this book. We are very grateful to Jonas Ohlsson, editor-in-chief at Nordicom, University of Gothenburg, Sweden, for making an Open Access publication possible, and to manuscript editor Kristin Clay, for her very helpful editorial suggestions and support. Additional thanks go to Matthieu Mercier for his graphic design skills and to all those who gave permission to use their tables and figures. The work to revise and update the report for this book received no financial support or input from the report's earlier contributors. We, as the book's authors, are fully responsible for any errors or omissions.

The Observatory is a core project of the Forum on Information and Democracy, the implementing civil society-led body of the Intergovernmental Partnership of the same name, endorsed by 55 democratic states. Its work aims to inform the international community's efforts to foster the adoption of effective and proportionate regulatory and non-regulatory measures for the protection of human rights – including the right to reliable information – and democracy in the digital space. The Observatory aims to provide a common and shared understanding of information ecosystems and their impact on democracy by aggregating and synthesising existing research through a robust, inclusive, and critical review process. By acting as a global research-to-policy interface in the field of information and democracy, the Observatory works to foster a more evidence-based roadmap towards both governmental and corporate accountability, and appropriate civic action to safeguard democracy. Our project was initiated by the late chair of the Forum on Information and Democracy, Christophe Deloire, to whom this book is

dedicated. A Prefiguration Group, co-chaired by Shoshana Zuboff and Ángel Gurriá (see Table A3 in the Appendix for affiliations and members) defined the objectives, methodology, and operational framework for our work.

The Observatory's first cycle report benefitted from two years of work by the Observatory's team, with input from a large number of contributors (see Table A3). For the work on the earlier report, we are enormously grateful for the strategic overview and project management support provided by the staff of the Observatory, led by Iris Boyer, then head of the Observatory, and Emma Gruden, project officer. We also extend our thanks to Michael Bak, former executive director of the Forum, and to Camille Grenier, current executive director, for their leadership, to Julie Pailhes for her support and especially her fundraising, as well as to Alice Boeffard-Dosierre, Chiara Gottardo, Giovanni Maggi, and Katharina Zuegel.

The work to prepare the first cycle report benefitted from a Steering Committee co-chaired by Dr. Courtney C. Radsch and Professor Gustavo Cardoso. Members are listed in Table A3, and we are hugely grateful for their support and input. We especially thank Luca Belli, Helani Galpaya, Jeanette Hofmann, Ansgar Koene, Nnenna Nwakanma, Pier Luigi Parcu, Courtney Radsch, and Natalia Zuazo for their sustained engagement and critical feedback throughout.

The Observatory coordinated three Research Assessment Panels, whose members are listed in Table A3. These experts helped enormously to aggregate and synthesise sources and commented on early drafts. We are grateful to all of them for their time and continuous engagement and, particularly, for very helpful comments and editing suggestions by Mina Aslam, Jens Berland, Beatriz Burarque, Jenny Fung, Giovanni De Gregorio, Poncelet Ilelej, Devika Mehta, Siddharta Menon, Sasha Meuter, Vincent Obia, Joao Peshanki, Pauline Renaud, Nikos Sarris, Anass Sedrati, Theresa Seip, Sharon Strover, Michela Sullivan-Paul, and Kate Wright. Text contributions by Anurag, Jeanette Hofmann, and Sharon Strover were incorporated into our text.

The report benefitted in addition from input from hundreds of academics, members of civil society organisations, and think tank representatives who took part in Stakeholder Advisory Group consultation meetings and representatives of states and public institutions of the Partnership for Information and Democracy. We warmly thank them all. We also received input from a large number of external contributors who shared their research in response to global calls for contributions. We thank them for their commitment, especially J. Khadijah Abdurahman, Natalia Aruguete, Jhalak Kakkar, Eugenia Mitchelstein, Syed Nazakat, Rasmus Nielsen, and Marianna Poyares, who agreed to be interviewed. Input was also gratefully received from think tanks, research institutes and other organisations, including the Institute for Strategic Dialogue (ISD), the Internet Freedom Foundation, IT

for Change, and UNESCO, as well as national regulators, with valuable insights provided by Ofcom observers, Jordan Ogg and Mark Pearson. The report benefited from Dawn Rushen's excellent copyediting skills, and the first report cycle was funded by the Postcode Foundation, the Ford Foundation, Luminate, and the French Ministry of Foreign Affairs.

Authors

Robin Mansell, professor emerita, Department of Media and Communications, London School of Economics and Political Science, United Kingdom

Flavia Durach, PhD, associate professor, head of department, Department of Digital Communication and Emerging Media, National University of Political Studies and Public Administration, Romania

Matthias C. Kettemann, professor and head of department, Department of Theory and Future of Law, University of Innsbruck, Austria; research program leader, Alexander von Humboldt Institute for Internet and Society and Leibniz Institute for Media Research, Hans-Bredow-Institute, Germany

Théophile Lenoir, PhD candidate, University of Milan, Italy; guest doctoral student, Sciences Po Medialab, France

Rob Procter, professor of social informatics, Department of Computer Science, University of Warwick; fellow, Alan Turing Institute for Data Science and AI, United Kingdom

Gyan Prakash Tripathi, lawyer and policy analyst, India; graduate scholar, Central European University, Austria

Emily Tucker, executive director, Center on Privacy & Technology at Georgetown Law, United States

Bibliography

- 2020 Election Research Project. (2020, August 31). A proposal for understanding social media's impact on elections: Peer-reviewed scientific research. *Medium*.
https://medium.com/@2020_election_research_project/a-proposal-for-understanding-social-medias-impact-on-elections-4ca5b7aae10
- Aalberg, T., Blekesaune, A., & Elvestad, E. (2013). Media choice and informed democracy: Toward increasing news consumption gaps in Europe? *The International Journal of Press/Politics*, 18(3), 281–303. <https://doi.org/10.1177/1940161213485990>
- Aaronson, S. A. (2021). Data is different, policymakers should pay attention to its governance. In M. Burri (Ed.), *Big data and global trade law* (pp. 340–360). Cambridge University Press.
- Aaronson, S. A., & Leblond, P. (2018). Another digital divide: The rise of data realms and its implications for the WTO. *Journal of International Economic Law*, 21(2), 245–272. <https://doi.org/10.1093/jiel/jgy019>
- Abdalla, M., & Abdalla, M. (2021). The grey hoodie project: Big tobacco, big tech, and the threat on academic integrity. *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. <https://doi.org/10.1145/3467102.3462563>
- Abdulrauf, L. A., & Dube, H. (Eds.). (2024). *Data privacy law in Africa: Emerging perspectives*. Pretoria University Law Press.
- Acemoglu, D., & Johnson, S. (2023). *Power and progress: Our thousand-year struggle over technology and prosperity*. Basic Books.
- Adams, R. (2022). *AI in Africa: Key concerns and policy considerations for the future of the continent* [Policy brief]. Africa Policy Research Institute (APRI), Germany.
<https://afripoli.org/ai-in-africa-key-concerns-and-policy-considerations-for-the-future-of-the-continent>
- Adel, K., Elhakeem, A., & Marzouk, M. (2023). Decentralized system for construction projects data management using blockchain and IPFS. *Journal of Civil and Engineering Management*, 29(4), 342–359. <https://doi.org/10.3846/jcem.2023.18646>
- Adjin-Tettey, T. D. (2022). Combating fake news, disinformation, and misinformation: Experimental evidence for media literacy education. *Cogent Arts & Humanities*, 9(1), 1–17. <https://doi.org/10.1080/23311983.2022.2037229>
- ADRN (African Digital Rights Network). (2024). *The digitalisation of social protection in Africa: The perspective of marginalised workers*.
<https://www.africandigitalrightsnetwork.org/publications>
- Advox Team. (2024). *The unfreedom monitor: A methodology for tracking digital authoritarianism around the world*. Global Voices Advox, DW Akademie and German Cooperation (Deutsche Zusammenarbeit).
https://globalvoices.org/wp-content/uploads/2023/08/Unfreedom_Monitor_Observatory_Final_Report_2023.pdf
- Afina, Y., Buchser, M., Krasodomski, A., Rowe, J., Sun, N., & Wilkinson, R. (2024). *Towards a global approach to digital platform regulation: Preserving openness amid the push for internet sovereignty*. Chatham House, Royal Institute of International Affairs.
- Africa Center for Strategic Studies. (2024). *Mapping a surge of disinformation in Africa*.
<https://africacenter.org/spotlight/mapping-a-surge-of-disinformation-in-africa>
- African Union. (2014). *African Union convention on cyber security and personal data protection*.
<https://africanlii.org/akn/aa-au/act/convention/2014/cyber-security-and-personal-data-protection/eng@2014-06-27>
- African Union. (2024, August 9). *Continental artificial intelligence strategy: Harnessing AI for Africa's development and prosperity*. https://au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf
- Aguerre, C., Campbell-Verduyn, M., & Scholte, J. A. (Eds.). (2024). *Global digital data governance: Polycentric perspectives*. Routledge. <https://doi.org/10.4324/9781003388418>
- Aharoni, T., Kligler-Vilenchik, N., & Tenenboim-Weinblatt, K. (2021). “Be less of a slave to the news”: A text-material perspective on news avoidance among young adults. *Journalism Studies*, 22(1), 42–59. <https://doi.org/10.1080/1461670X.2020.1852885>
- Ahmad, S., & Greb, M. (2022). Automating social media content moderation: Implications for governance and labour discretion. *Work in the Global Economy*, 2(2), 176–198.
<https://doi.org/10.1332/273241721X16647876031174>

- Ahmed, N., Wahed, M., & Thompson, N. C. (2023). The growing influence of industry in AI research. *Science*, 379(6635), 884–886. <https://doi.org/10.1126/science.ade2420>
- AI Elections Accord. (2024). *A tech accord to combat deceptive use of AI in 2024 elections*. <https://securityconference.org/en/aielectionsaccord/>
- AI Now Institute. (2024). *Redirecting Europe's AI industrial policy: From competitiveness to public interest*. <https://ainowinstitute.org/redirecting-europes-ai-industrial-policy>
- Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: A review. *Social Network Analysis and Mining*, 13(30), 1–36. <https://doi.org/10.1007/s13278-023-01028-5>
- Ajuntament de Barcelona [Barcelona City Council]. (2023). *Definició de metodologies de treball i protocols per a la implementació de sistemes algorítmics [Definition of work methodologies and protocols for the implementation of algorithmic systems]*. https://ajuntament.barcelona.cat/digital/sites/default/files/2023-10/def_metodologia_sist_algoritmics_cat.pdf
- Ajunwa, I., Crawford, K., & Schultz, J. (2017). Limitless worker surveillance. *California Law Review*, 105(3), 735–776. <https://www.jstor.org/stable/44630759>
- Akimenko, V., & Giles, K. (2020). Russia's cyber and information warfare. *Asia Policy*, 15(2), 67–75. <https://www.jstor.org/stable/27023902>
- Akman, P. (2022). Regulating competition in digital platform markets: A critical assessment of the framework and approach of the EU Digital Markets Act. *European Law Review*, 47(2022), 85–114. <https://dialnet.unirioja.es/servlet/articulo?codigo=8320879>
- Akser, M., & Baybars, B. (2023). Repressed media and illiberal politics in Turkey: The persistence of fear. *Southeast European and Black Sea Studies*, 23(1), 159–177. <https://doi.org/10.1080/14683857.2022.2088647>
- Alaimo, C., & Kallinikos, J. (2024). *Data rules: Reinventing the market economy*. MIT Press. <https://doi.org/10.7551/mitpress/11751.001.0001>
- Albaladi, A., Islam, M., Das, A., Bigonah, M., Zhang, Z., Jamshidi, F., Rahgouy, M., Raychawdhary, N., Marghitu, D., & Seals, C. (2025). Hate speech detection using large language models: A comprehensive review. IEEE Access. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10848067>
- Alexander, F. (2024, July 24). *UN attempts AI power grab: The West is unhappy*. CEPA. <https://cepa.org/article/un-attempts-ai-power-grab-the-west-is-unhappy>
- Alfonsi, C. (2019). Taming tech giants requires fixing the revolving door. *Harvard Kennedy School Review*, 19(2019), 166–170. <https://www.proquest.com/docview/2316725199/fulltextPDF/4F677B3FE2CA46B1PQ/32>
- Ali S. M., Paragi, B., Daly, A. C., Gjorgjioska, A., Hespanhol, L., Kerasidou, X., Kouadri Mostéfaoui, S., Oyeniji, O., & Tomičić, A. (2024). The (un)bearable whiteness of AI ethics. In D. J. Gunkel (Ed.), *Handbook on the ethics of artificial intelligence* (pp. 218–231). Edward Elgar Publishing. <https://doi.org/10.4337/9781803926728.00020>
- Alieva, I., Moffitt, J. D., & Carley, K. M. (2022). How disinformation operations against Russian opposition leader Alexei Navalny influence the international audience on Twitter. *Social Network Analysis and Mining*, 12(80), 1–13. <https://doi.org/10.1007/s13278-022-00908-6>
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 32(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Allen, D., & Weyl, E. G. (2024). The real dangers of generative AI. *Journal of Democracy*, 35(1), 146–162. <https://www.journalofdemocracy.org/articles/the-real-dangers-of-generative-ai>
- Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), 1–6. <https://doi.org/10.1126/sciadv.aay3539>
- Altay, S., & Acerbi, A. (2023). People believe misinformation is a threat because they assume others are gullible. *New Media & Society*, 26(11), 6440–6461. <https://doi.org/10.1177/14614448231153379>
- Altay, S., Berriche, M., & Acerbi, A. (2023a). Misinformation on misinformation: Conceptual and methodological challenges. *Social Media + Society*, 9(1), 1–13. <https://journals.sagepub.com/doi/epub/10.1177/20563051221150412>
- Altay, S., Fletcher, R., & Nielsen, R. K. (2025). News participation is declining: Evidence from 46 countries between 2015 and 2022. *New Media & Society*, 27(9), 5110–5131. <https://doi.org/10.1177/14614448241247822>

- Altay, S., Nielsen, R. K., & Fletcher, R. (2023b). News can help! The impact of news media and digital platforms on awareness of and belief in misinformation. *The International Journal of Press/Politics*, 29(2), 459–484. <https://doi.org/10.1177/19401612221148981>
- Alufaisan, Y., Marusich, L. R., Bakdash, J. Z., Zhou, Y., & Kantarcioglu, M. (2021). Does explainable artificial intelligence improve human decision-making? *Proceeding of the AAAI Conference on Artificial Intelligence*, 35(8), 6618–6626. <https://doi.org/10.1609/aaai.v35i8.16819>
- American Sunlight Project. (2024). *The information laundering cycle: How a coordinated effort weaponized the American political system in favor of disinformation*. <https://www.americansunlight.org/ilc-database>
- Amnesty International & AI Forensics. (2023). *Driven into the darkness: How TikTok's 'for you' feed encourages self-harm and suicidal ideation*. <https://www.amnesty.org/en/documents/POL40/7350/2023/en>
- Ananny, M. (2023). What a “platform press” view has to offer. *Digital Journalism*, 11(8), 1568–1575. <https://doi.org/10.1080/21670811.2023.2257768>
- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989. <https://doi.org/10.1177/1461444816676645>
- Andere, B., & Kathure, M. (2024). *Strengthening data protection in Africa: Key issues for implementation*. Access Now. <https://www.accessnow.org/wp-content/uploads/2024/01/Strengthening-data-protection-in-Africa-key-issues-for-implementation-updated.pdf>
- Anderson, B. (1983). *Imagined communities: Reflections on the origin and spread of nationalism*. Knopf Doubleday.
- Anderson, C. W. (2021). Fake news is not a virus: On platforms and their effects. *Communication Theory*, 31(1), 42–61. <https://doi.org/10.1093/ct/qtaa008>
- Anderson, N. (2024, August 23). *After cybersecurity lab wouldn't use AV software, US accuses Georgia Tech of fraud*. Ars Technica. <https://arstechnica.com/security/2024/08/oh-your-cybersecurity-researchers-wont-use-antivirus-tools-heres-a-federal-lawsuit>
- Ang, P.-H., & Haristya, S. (2024). The governance, legitimacy and efficacy of Facebook's oversight board: A model for global tech platforms? *Emerging Media*, 2(2), 169–180. <https://doi.org/10.1177/27523543241266860>
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). *Machine bias*. ProPublica. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Annan, K. (1997, June 23). *If information and knowledge are central to democracy, they are conditions for development* [Press release]. United Nations <https://press.un.org/en/1997/19970623.sgsm6268.html>
- Annoni, A., Benczur, P., Bertoldi, P., Delipetrev, B., De Prato, G., Feijoo, C., Fernandez Macias, E., Gomez Gutierrez, E., Iglesias Portela, M., Junklewitz, H., Lopez Cobo, M., Martens, B., Figueiredo Do Nascimento, S., Nativi, S., Polvora, A., Sanchez Martin, J. I., Tolan, S., Tuomi, I., & Vesnic Alujevic, L. (2018). *Artificial intelligence: A European perspective*. Publications Office of the European Union. <https://doi.org/10.2760/11251>
- Anstead, N., Edwards, L., Livingstone, S., & Stoilova, M. (2025). The potential for media literacy to combat misinformation: Results of a rapid evidence assessment. *International Journal of Communication*, 19(2025), 2129–2151. <https://ijoc.org/index.php/ijoc/article/view/23549>
- Anthony, A., Sharma, L., & Noor, E. (2024, April 30). *Advancing a more global agenda for trustworthy artificial intelligence*. Carnegie Endowment for International Peace. <https://carnegieendowment.org/research/2024/04/advancing-a-more-global-agenda-for-trustworthy-artificial-intelligence?lang=en>
- Arcom. (2024). *Les Français et information [The French and information]*. <https://www.arcom.fr/nos-ressources/etudes-et-donnees/etudes-bilans-et-rapports-de-larcom/les-francais-et-linformation>
- Arendt, H. (1998). *The human condition*. University of Chicago Press. (Original work published 1958)
- Arendt, H. (1968). Truth and politics. In H. Arendt, *Between past and future: Eight exercises in political thought* (pp. 227–264). Viking Press.
- Arendt, H. (1978). Hannah Arendt: From an interview. *The New York Review of Books*, 25(16). <https://www.nybooks.com/articles/1978/10/26/hannah-arendt-from-an-interview/>

- Arriagada, A., Bonhomme, M., Ibanez, F., & Leyton, J. (2023). The gig economy in Chile: Examining labor conditions and the nature of gig work in a Global South country. *Digital Geography and Society*, 5(2023), 1–6. <https://doi.org/10.1016/j.diggeo.2023.100063>
- Aruguete, N., & Calvo, E. (2023). *Nosotros contra ellos: Cómo trabajan las redes para confirmar nuestras creencias y rechazar las de los otros* [Us against them: How networks work to confirm our beliefs and reject those of others]. *Sociologie y Política*.
- Ashok, P. (2023). The curious case of automated decision-making in India. *International Cybersecurity Law Review*, 4(2), 235–248. <https://doi.org/10.1365/s43439-022-00078-3>
- Aufderheide, P. (1993). *Media literacy: A report of the national leadership conference on media literacy*. Aspen Institute Communications and Society Program. <https://eric.ed.gov/?id=ED365294>
- Aula, V. (2023). Evidence-based policymaking in the legislatures: Timeliness and politics of evidence in Finland. *Policy & Politics*, 51(4), 673–694. <https://doi.org/10.1332/030557321X16873376680333>
- Avle, S. (2022). Hardware and data in the platform era: Chinese smartphones in Africa. *Media, Culture & Society*, 44(8), 1473–1489. <https://doi.org/10.1177/01634437221128935>
- AWS (Amazon Web Services). (2025). *Amazon just walk out*. <https://aws.amazon.com/just-walk-out>
- Bachelet, M. (2019). *Human rights in the digital age – Can they make a difference?* OHCHR Keynote Speech. <https://www.ohchr.org/en/speeches/2019/10/human-rights-digital-age>
- Baecker, C., Alabbadi, O., Yogiputra, G. P., & Tien Dung, N. (2023). Threats provided by artificial intelligence that could disrupt the democratic system [Scientific paper, University of Applied Science, Brandenburg, Germany].
- Bagchi, C., Menczer, F., Lundquist, J., Tarafdar, M., Paik, A., & Grabowicz, P. (2024). Social media algorithms can curb misinformation, but do they? *ArXiv*. <https://doi.org/10.48550/arXiv.2409.18393>
- Bail, C. (2021). *Breaking the social media prism: How to make our platforms less polarizing*. Princeton University Press.
- Bailey, R., Parsheera, S., & Sane, R. (2020). *Comments on the ‘report by the committee of experts on non-personal data governance framework’ to Ministry of Electronics and IT, Government of India*. <https://doi.org/10.2139/ssrn.3724184>
- Baines, D., & Elliott, R. (2020). *Defining misinformation, disinformation and malinformation: An urgent need for clarity during the COVID-19 infodemic* [Discussion papers 20-16, Department of Economics, University of Birmingham, England]. <https://ideas.repec.org/p/bir/birmec/20-06.html>
- Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital Journalism*, 6(2), 154–175. <https://doi.org/10.1080/21670811.2017.1345645>
- Bakir, V., & McStay, A. (2022). Core incubators of false information online. In V. Bakir, & A. McStay, *Optimising emotions, incubating falsehoods: How to protect the global civic body from disinformation and misinformation* (pp. 29–52). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-13551-4_2
- Bakke, N. A., & Barland, J. (2022). Disruptive innovations and paradigm shifts in journalism as a business: From advertisers first to readers first and traditional operational models to the AI factory. *SAGE Open*, 12(2), 1–18. <https://doi.org/10.1177/21582440221094819>
- Bakker, B. N., & Lelkes, Y. (2024). Putting the affect into affective polarisation. *Cognition and Emotion*, 38(4), 418–436. <https://doi.org/10.1080/02699931.2024.2362366>
- Balkin, J. M. (2017). Free speech in the algorithmic society: Big data, private governance, and new school speech regulation. *UC Davis Law Review*, 51, 1151–1209. <https://lawreview.law.ucdavis.edu/archives/51/3/free-speech-algorithmic-society-big-data-private-governance-and-new-school-speech>
- Ball, K., & Snider, L. (2013). *The surveillance-industrial complex: A political economy of surveillance*. Routledge. <https://doi.org/10.4324/9780203094426>
- Banaji, S. (2024). Antidigital resilience: The (anti-)ethics of participation in an unjust and unequal public sphere. *Javnost – The Public*, 31(1), 141–157. <https://doi.org/10.1080/13183222.2024.2311012>

- Banaji, S., & Bhat, R. (2022). *Social media and hate*. Routledge.
<https://doi.org/10.4324/9781003083078>
- Banaji, S., Bhat, R., Agarwal, A., Passanha, N., & Pravin, M. S. (2019). *WhatsApp vigilantes: An exploration of citizen reception and circulation of WhatsApp misinformation linked to mob violence in India*. Department of Media and Communications, London School of Economics and Political Science. https://eprints.lse.ac.uk/104316/1/Banaji_whatsapp_vigilantes_exploration_of_citizen_reception_published.pdf
- Banerjee, S., Mont'Alverne, C., Ross Arguedas, A., Toff, B., Fletcher, R., & Nielsen, R. K. (2023). *Strategies for building trust in news: What the public say they want across four countries* Reuters Institute for the Study of Journalism, University of Oxford.
<https://ora.ox.ac.uk/objects/uuid:689e057a-bfad-4b3d-a69c-7e83eac8ff1>
- Bannerman, S. (2022). Platform imperialism, communications law and relational sovereignty. *New Media & Society*, 26(4), 1816–1833. <https://doi.org/10.1177/14614448221077284>
- Bannerman, S., McNeil, B., Wyndham-West, K., & Beny, H. (2024). *2023 TechLobby annual report*. The Tech Lobby.
<https://thetechlobby.ca/2023-techlobby-annual-report-interactive-graphics>
- Bansal, V. (2024, April 12). *This delivery app takes away health insurance when workers don't meet quotas*. Rest of World Reporting Global Tech Stories.
<https://restofworld.org/2024/swiggy-health-insurance-quotas>
- Baptista, J. P., & Gradim, A. (2022). Online disinformation on Facebook: The spread of fake news during the Portuguese 2019 election. *Journal of Contemporary European Studies*, 30(2), 297–312. <https://doi.org/10.1080/14782804.2020.1843415>
- Baranes, E. (2014). The interplay between network investment and content quality: Implications to net neutrality on the Internet. *Information Economics and Policy*, 28, 57–69.
<https://doi.org/10.1016/j.infoecopol.2014.07.002>
- Barba-Kay, A. (2023). *A web of our own making: The nature of digital formation*. Columbia University Press. <https://doi.org/10.1017/9781009324786>
- Bargaoanu, A., & Radu, L. (2018). Fake news or disinformation 2.0? Some insights into Romanians' digital behaviour. *Romanian Journal of European Affairs*, 18(1), 24–38.
<https://ssrn.com/abstract=3296922>
- Barocas, S., & Nissenbaum, H. (2014a). Big data's end run around anonymity and consent. In J. Lane, V. Stodden, S. Bender, & H. Nissenbaum (Eds.), *Privacy, big data, and the public good: Frameworks for engagement* (pp. 44–75). Cambridge University Press.
<https://doi.org/10.1017/CBO9781107590205.004>
- Barocas, S., & Nissenbaum, H. (2014b). Big data's end run around procedural privacy protections. *Communications of the ACM*, 57, 31–33.
<https://cacm.acm.org/opinion/big-datas-end-run-around-procedural-privacy-protections/>
- Baron, G. R. (2002). Now is too late: Survival in an era of instant news. *Financial Times*. Prentice Hall.
- Bateman, J., & Jackson, D. (2024). *Looking ahead: Generative AI – Countering disinformation effectively: An evidence-based policy guide*. Carnegie Endowment for International Peace.
<https://carnegieendowment.org/2024/01/31/looking-ahead-generative-ai-pub-91489>
- Bates, J., Kennedy, H., Perea, I. M., Oman, S., & Pinney, L. (2023). Socially meaningful transparency in data-based systems: Reflections and proposals from practice. *Journal of Documentation*, 80(1), 1–20. <https://doi.org/10.1108/jd-01-2023-0006>
- Bauer, J. M., & Knieps, G. (2018). Complementary innovation and network neutrality. *Telecommunications Policy*, 42(2), 172–183. <https://doi.org/10.1016/j.telpol.2017.11.006>
- Bayer, J., Katsirea, I., Batura, O., Holznagel, B., Hartmann, S., & Lubianiec, K. (2021). *The fight against disinformation and the right to freedom of expression*. Report for the European Parliament's Committee on Civil Liberties, Justice and Home Affairs. Universities of Münster, Sheffield, Leuphana, and Ecorys, Poland. [https://www.europarl.europa.eu/RegData/etudes/STUD/2021/695445/IPOL_STU\(2021\)695445_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2021/695445/IPOL_STU(2021)695445_EN.pdf)
- BBC Media Action. (2021). *Enabling media markets to work for democracy: An international fund for public interest media*. International Fund for Public Interest Media (IFPIM).
<https://ifpim.org/resources/feasibility-study/>

- Beckett, C. (2019). *New powers, new responsibilities: A global survey of journalism and artificial intelligence*. POLIS (Department of Media and Communications), London School of Economics and Political Science; Google News Initiative.
<https://www.journalismai.info/research/2019-new-powers-new-responsibilities>
- Beckett, C., & Yaseen, M. (2023). *Generating change: A global survey of what news organisations are doing with AI*. POLIS (Department of Media and Communications), London School of Economics and Political Science; Google News Initiative.
<https://www.journalismai.info/research/2023-generating-change>
- Beiguelman, G. (2021). *Políticas da imagem: vigilância e resistência na dadosfera [Image politics: Surveillance and resistance in the datasphere]*. Ubu Editora.
- Belenguer, L. (2022). AI bias: Exploring discriminatory algorithmic decision-making models and the application of possible machine-centric solutions adapted from the pharmaceutical industry. *AI and Ethics*, 2(4), 771–787. <https://doi.org/10.1007/s43681-022-00138-8>
- Bell, E., Owen, T., Brown, P. D., Hauka, C., & Rashidian, N. (2017). *The platform press: How Silicon Valley reengineered journalism*. Tow Center for Journalism, Columbia Journalism School. <https://www.mediatechdemocracy.com/all-work/the-platform-press>
- Bellan, R. (2024, August 15). Meta axed CrowdTangle, a tool for tracking disinformation. Critics claim its replacement has just ‘1% of the features’. *TechCrunch*.
<https://techcrunch.com/2024/08/15/meta-shut-down-crowdtangle-a-tool-for-tracking-disinformation-heres-how-its-replacement-compares>
- Belli, L., Almeida, C., Couto, N., Radu, R., Weber, R. H., & Brown, I. (2022). *Towards meaningful and interoperable transparency for digital platform: 2022 outcome of the UN IGF coalition on platform responsibility*. Internet Governance Forum.
https://www.intgovforum.org/en/filedepot_download/57/23886
- Bello y Villarino, J.-M. (2023). Global standard-setting for artificial intelligence: Para-regulating international law for AI? *The Australian Year Book of International Law Online*, 41(1), 57–181. <https://doi.org/10.1163/26660229-04101018>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623.
<https://doi.org/10.1145/3442188.3445922>
- Benequista, N., Abbott, S., Rothman, P., & Mano, W. (Eds.). (2019). *International media development: Historical perspectives and new frontiers*. Peter Lang.
- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. John Wiley & Sons.
- Benkler, Y. (2020). A political economy of the origins of asymmetric propaganda in American media. In S. Livingston, & W. L. Bennett (Eds.), *The disinformation age* (pp. 43–66). Cambridge University Press. <https://doi.org/10.1017/9781108914628>
- Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.
<https://doi.org/10.1093/oso/9780190923624.001.0001>
- Bennett, C. J., & Raab, C. (2006). *The governance of privacy: policy instruments in a global perspective*. MIT Press.
- Bennett, W. L., & Kneuer, M. (2023). Communication and democratic erosion: The rise of illiberal public spheres. *European Journal of Communication*, 39(2), 1–20.
<https://doi.org/10.1177/02673231231217378>
- Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139. <https://doi.org/10.1177/0267323118760317>
- Bennett, W. L., & Livingston, S. (2023). A brief history of the disinformation age: Information wars and the decline of institutional authority. In S. Salgado, & S. Papathanassopoulos (Eds.), *Streamlining political communication concepts: Updates, changes, normalcies* (pp. 43–73). Springer. <https://doi.org/10.1007/978-3-031-45335-9>
- Bennett, W. L., & Livingston, S. (Eds.). (2020). *The disinformation age: Politics, technology and disruptive communication in the United States*. Cambridge University Press.
<https://doi.org/10.1017/9781108914628>

- Benson, J. (2019). Deliberative democracy and the problem of tacit knowledge. *Politics, Philosophy & Economics*, 18(1), 76–97. <https://doi.org/10.1177/1470594X18782086>
- Benton, J. (2022, June 16). *Facebook looks ready to divorce the news industry, and I doubt couples counseling will help*. NiemanLab. <https://www.niemanlab.org/2022/06/facebook-looks-ready-to-divorce-the-news-industry-and-i-doubt-couples-counseling-will-help>
- Berdou, E. (2011). *Organization in open source communities: At the crossroads of the gift and market economies*. Routledge. <https://doi.org/10.4324/9780203851975>
- BEREC (Body of European Regulators for Electronic Communications). (2024a). *BEREC guidelines on the implementation by national regulators of European net neutrality rules*. <https://www.berec.europa.eu/en/document-categories/berec/regulatory-best-practices/guidelines/berec-guidelines-on-the-implementation-by-national-regulators-of-european-net-neutrality-rules>
- BEREC (Body of European Regulators for Electronic Communications). (2024b). *What is zero-rating?* <https://www.berec.europa.eu/en/what-is-zero-rating>
- Berger, C., Freihse, C., Kettemann, M. C., Mosene, K., & Hofmann, V. (2023a, August 28). *The EU elections 2024: How to build resilience against disinformation campaigns on social platforms*. Upgrade Democracy. <https://upgradedemocracy.de/en/impulse/the-eu-elections-2024-how-to-build-resilience-against-disinformation-campaigns-on-social-platforms>
- Berger, C., Freihse, C., Mosene, K., Kettemann, M. C., & Hofmann, V. (2023b, July 26). *Threats to democracy: Climate misinformation and gendered disinformation*. Upgrade Democracy. <https://upgradedemocracy.de/en/impulse/climate-misinformation-and-gender-related-disinformation-responsibilities-of-civil-society-public-sector-and-the-media>
- Berger, G., Gillwald, A., Orembo, E., Diouf, D., & Garcia, J. M. (2023). *Platform problems and regulatory solutions: Findings from a comprehensive review of existing studies and investigations*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000385813>
- Bernardi, A. J. B., & de Morais, J. A. (2021). Fascismo à brasileira? Análise de conteúdo dos discursos de Bolsonaro após o segundo turno das eleições presidenciais de 2018 [Brazilian fascism? Content analysis of Bolsonaro's speeches after the second round of the 2018 presidential elections]. *Política & Sociedade*, 20(48), 300–327. <https://doi.org/10.5007/2175-7984.2021.72401>
- Berthelot, A., Caron, E., Jay, M., & Lefèvre, L. (2024). Estimating the environmental impact of Generative-AI services using an LCA-based methodology. *31st CIRP Conference on Life Cycle Engineering*, 122, 707–712. <https://doi.org/10.1016/j.procir.2024.01.098>
- Beskow, D. M., & Carley, K. M. (2020). Characterization and comparison of Russian and Chinese disinformation campaigns. In K. Shu, S. Wang, D. Lee, & H. Liu (Eds.), *Disinformation, misinformation, and fake news in social media: Emerging research challenges and opportunities* (pp. 63–81). Springer. https://doi.org/10.1007/978-3-030-42699-6_4
- Besson, S. (2011). Human rights and democracy in a global context: Decoupling and recoupling. *Ethics & Global Politics*, 4(1), 19–50. <https://doi.org/10.3402/egp.v4i1.6348>
- Bhat, P. (2021). Platform politics: The emergence of alternative social media in India. *Asia Pacific Media Educator*, 31(2), 269–276. <https://doi.org/10.1177/1326365X211056699>
- Bhat, P., & Chadha, K. (2020). Anti-media populism: Expressions of media distrust by right-wing media in India. *Journal of International and Intercultural Communication*, 13(2), 166–182. <https://doi.org/10.1080/17513057.2020.1739320>
- Bhatia, K. V., Elhussein, M., Kreimer, B., & Snapp, T. (2023). Protests, internet shutdowns, and disinformation in a transitioning state. *Media, Culture & Society*, 45(6), 1101–1118. <https://doi.org/10.1177/01634437231155568>
- Biddle, S. (2024, March 25). *Elon Musk fought government surveillance – While profiting off government surveillance*. *The Intercept*. <https://theintercept.com/2024/03/25/elon-musk-x-dataminr-surveillance-privacy>
- Biju, P. R., & Gayathri, O. (2023). Self-breeding fake news: Bots and artificial intelligence perpetuate social polarization in India's conflict zones. *The International Journal of Information, Diversity, & Inclusion*, 7(1/2), 1–25. <https://doi.org/10.33137/ijidi.v7i1/2.39409>
- Bimber, B., & Gil de Zúñiga, H. (2020). The unedited public sphere. *New Media & Society*, 22(4), 700–715. <https://doi.org/10.1177/1461444819893980>

- Bineham, J. L. (1988). A historical account of the hypodermic model in mass communication. *Communication Monographs*, 55(3), 230–246. <https://doi.org/10.1080/03637758809376169>
- Birkinbine, B. J. (2018). Commons praxis: Towards a critical political economy of the digital commons. *tripleC*, 16(1), 290–305. <https://doi.org/10.31269/tripleC.v16i1.929>
- Birnbaum, E., & Martin, E. (2023, July 10). Critics slam big tech lobbying in US Indo-Pacific talks. *TIME*. <https://time.com/6293489/tech-firms-us-indo-pacific-trade-talks>
- Birwe, H. (2024). *La désinformation basée sur le genre en contexte de crises: Quels mécanismes de prévention et dispositifs de lutte pour les pays du Sahel? [Gender-based disinformation in the context of crises: What prevention mechanisms and combat measures are available for Sahel countries?]*. IDOS Discussion Paper, Géographe diplômé de Sciences Po Lyon et de l'Université Cheikh Anta Diop de Dakar. <https://www.idos-research.de/discussion-paper/article/information-integrity-and-information-pollution-vulnerabilities-and-impact-on-social-cohesion-and-democracy-in-mexico/>
- Bitter, A. (2024, August 3). Amazon's just walk out: Technology relies on hundreds of workers in India watching you shop. *Business Insider India*. <https://www.businessinsider.in/retail/news/amazons-just-walk-out-technology-relies-on-hundreds-of-workers-in-india-watching-you-shop/articleshow/109014034.cms>
- Blake, E., Odendaal, N., & Söderström, O. (2023). Beyond 'data positivism': Civil society organizations' data and knowledge tactics in South Africa. In O. Söderström, & A. Datta (Eds.), *Data power in action* (pp. 204–226). Bristol University Press.
- Blasi Casagran, C., & Vermeulen, M. (2021). Reflections on the murky legal practices of political micro-targeting from a GDPR perspective. *International Data Privacy Law*, 11(4), 348–359. <https://doi.org/10.1093/idpl/ipab018>
- Blenkinsop, P. (2025, January 8). *We do not censor social media, EU says in response to Meta*. Reuters. <https://www.reuters.com/technology/we-do-not-censor-social-media-eu-says-response-meta-2025-01-08/>
- Blöbaum, B. (Ed.). (2016). *Trust and communication in a digitized world: Models and concepts of trust research*. Springer. <https://doi.org/10.1007/978-3-319-28059-2>
- Blum, I., & Uldam, J. (2024). Faking, optimising and conceding to power: Social movement understandings of social media power. *New Media & Society*, Online First, 1–19. <https://doi.org/10.1177/14614448241266769>
- Boadle, A. (2023, May 3). *Brazil pushes back on big tech firms' campaign against 'fake news law'*. Reuters. <https://www.reuters.com/world/americas/brazil-lawmakers-vote-controversial-bill-clean-up-social-media-2023-05-02>
- Boas, T. C. (2006). Weaving the authoritarian web: The control of internet use in nondemocratic regimes. In J. Zysman, & A. Newman (Eds.), *How revolutionary was the digital revolution? National responses, market transitions and global technology* (pp. 361–378). Stanford University Press.
- Boczkowski, P. J., & Mitchelstein, E. (2021). *The digital environment: How we live, learn, work, and play now*. MIT Press. <https://doi.org/10.7551/mitpress/13602.001.0001>
- Boczkowski, P. J., & Mitchelstein, E. (2022). *El Entorno Digital: Breve manual para entender cómo vivimos, aprendemos, trabajamos y pasamos el tiempo libre hoy [The digital environment: A brief handbook for understanding how we live, learn, work, and spend our free time today]*. Siglo Veintiuno.
- Boix, C. (2022). AI and the economic and informational foundations of democracy. In J. B. Bullock, Y.-C. Chen, J. Himmelreich, V. M. Hudson, A. Korinek, M. M. Young, & B. Zhang (Eds.), *The Oxford handbook of AI governance* (pp. 707–725). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197579329.001.0001>
- Bolin, G., & Kunelius, R. (2023). The return of propaganda: Historical legacies and contemporary conceptualisations. *Nordic Journal of Media Studies*, 5(1), 1–16. <https://doi.org/10.2478/njms-2023-0001>
- Bonfanti, M. E. (2020). *The weaponisation of synthetic media: What threat does this pose to national security?* Center for Security Studies (CSS), Zürich, Elcano Royal Institute for International and Strategic Studies (Real Instituto Elcano).
- Bonini, T., & Treré, E. (2024). *Algorithms of resistance: The everyday fight against platform power*. MIT Press.

- Bontcheva, K., Papadopoulous, S., Tsalakanidou, F., Gallotti, R., Dutkiewicz, L., Krack, N., Teyssou, D., Nucci, F. S., Spangenberg, J., Srba, I., Aichroth, P., Cuccovillo, L., & Verdoliva, L. (2024). *Generative AI and disinformation: Recent advances, challenges, and opportunities*. Vera.ai; AI4Trust; AI4Media; TITAN. https://edmo.eu/wp-content/uploads/2023/12/Generative-AI-and-Disinformation_-White-Paper-v8.pdf
- Bontcheva, K., Posetti, J., Teyssou, D., Meyer, T., Gregory, S., Hanot, C., & Maynard, D. (2020). *Balancing act: Countering digital disinformation while respecting freedom of expression*. Broadband Commission; ITU; UNESCO. <https://www.broadbandcommission.org/publication/balancing-act-countering-digital-disinformation>
- Bontemps, V. (2025, May 5). Meta content moderator cuts over 2,000 jobs in Spain: Union. Agence France-Presse. <https://www.france24.com/en/live-news/20250505-meta-content-moderator-cuts-over-2-000-jobs-in-spain-union>
- Bontridder, N., & Pouillet, Y. (2021). The role of artificial intelligence in disinformation. *Data & Policy*, 3(e32), 1–21. <https://doi.org/10.1017/dap.2021.20>
- Boots, B. C., Matlack, A. K., & Richardson-Gool, T. S. (2024). A call for promoting algorithmic literacy. *Proceedings of the TPRC2024, the Research Conference on Communications, Information and Internet Policy*. <https://doi.org/https://papers.ssrn.com/abstract=4912427>
- Borelli, M. (2023). Social media corporations as actors of counter-terrorism. *New Media & Society*, 25(11), 2877–2897. <https://doi.org/10.1177/14614448211035121>
- Borges, D., & Reviglio, U. (2024, March 27). Diversification is good: Emerging business models for local news media. EUI (Centre for Media Pluralism and Media Freedom). <https://cmpf.eui.eu/emerging-business-models-for-local-media>
- Borgogno, O., & Zangrandi, M. S. (2024). Chinese data governance and trade policy: From cyber sovereignty to the quest for digital hegemony? *Journal of Contemporary China*, 33(148), 578–602. <https://doi.org/10.1080/10670564.2023.2299961>
- Borocas, S., Hardt, M., & Narayanan, A. (2023). *Fairness and machine learning: Limitations and opportunities*. MIT Press.
- Borz, G., De Francesco, F., Montgomerie, T. L., & Bellis, M. P. (2024). The EU soft regulation of digital campaigning: Regulatory effectiveness through platform compliance to the code of practice on disinformation. *Policy Studies*, 45(5), 709–729. <https://doi.org/10.1080/01442872.2024.2302448>
- Bos, L., Kruijemeier, S., & de Vreese, C. (2016). Nation binding: How public service broadcasting mitigates political selective exposure. *PLOS ONE*, 11(5), 1–11. <https://doi.org/10.1371/journal.pone.0155112>
- Botta, M. (2021). Sector regulation of digital platforms in Europe: Uno, Nessuno e Centomila [One, no one, and one hundred thousand]. *Journal of European Competition Law & Practice*, 12(7), 500–512. <https://doi.org/10.1093/jeclap/lpab046>
- Boudry, C., Alvarez-Muñoz, P., Arencibia-Jorge, R., Ayena, D., Brouwer, N. J., Chaudhuri, Z., Chawner, B., Epee, E., Errais, K., Fotouhi, A., Gharaibeh, A. M., Hassanein, D. H., Herwig-Carl, M. C., Howard, K., Kaimbo, D. K. W., Laughrea, P.-A., Lopez, F. A., Machin-Mastromatteo, J. D., Malerbi, F. K. ... Mouriaux, F. (2019). Worldwide inequality in access to full text scientific articles: The example of ophthalmology. *PeerJ*, 7, 1–18. <https://doi.org/10.7717/peerj.7850>
- Boukes, M., & Vliegenthart, R. (2017). News consumption and its unpleasant side effect: Studying the effect of hard and soft news exposure on mental well-being over time. *Journal of Media Psychology*, 29(3), 137–147. <https://doi.org/10.1027/1864-1105/a000224>
- Boulianne, S., Koc-Michalska, K., & Bimber, B. (2020). Right-wing populism, social media and echo chambers in Western democracies. *New Media & Society*, 22(4), 683–699. <https://doi.org/10.1177/1461444819893983>
- Boulianne, S., Tenove, C., & Buffie, J. (2022). Complicating the resilience model: A four-country study about misinformation. *Media and Communication*, 10(3), 169–182. <https://doi.org/10.17645/mac.v10i3.5346>
- Bover, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications*, 10(1), 1–14. <https://doi.org/10.1038/s41467-018-07761-2>

- boyd, d. (2014). *It's complicated: The social lives of networked teens*. Yale University Press.
- Bradford, A. (2020). *The Brussels effect: How the European Union rules the world*. Oxford University Press. <https://doi.org/10.1093/oso/9780190088583.001.0001>
- Bradshaw, S., & DeNardis, L. (2022). *Internet infrastructure as an emerging terrain of disinformation*. CIGI (Centre for International Governance Innovation), Canada. <https://www.cigionline.org/articles/internet-infrastructure-as-an-emerging-terrain-of-disinformation>
- Bradshaw, S., & Howard, P. N. (2019). *The global disinformation order: 2019 global inventory of organised social media manipulation*. Oxford Internet Institute, University of Oxford. <https://www.oii.ox.ac.uk/news-events/reports/the-global-disinformation-order-2019-global-inventory-of-organised-social-media-manipulation>
- Bradshaw, S., Bailey, H., & Howard, P. N. (2021). *Industrialized disinformation: 2020 global inventory of organized social media manipulation*. Oxford Internet Institute, University of Oxford. <https://demtech.oii.ox.ac.uk/research/posts/industrialized-disinformation/>
- Bradshaw, S., Campbell-Smith, U., Henle, A., Perini, A., Shalev, S., Bailey, H., & Howard, P. N. (2020). *Country case studies, industrialized disinformation: 2020 global inventory of organized social media manipulation*. Oxford Internet Institute, University of Oxford. https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/03/Case-Studies_FINAL.pdf
- Brandtzaeg, P. B., Følstad, A., & Chaparro Domínguez, M. Á. (2018). How journalists and social media users perceive online fact-checking and verification services. *Journalism Practice*, 12(9), 1109–1129. <https://doi.org/10.1080/17512786.2017.1363657>
- Brayne, S. (2020). *Predict and surveil: Data, discretion, and the future of policing*. Oxford University Press.
- Breuer, A. (2024). *Information integrity and information pollution: Vulnerabilities and impact on social cohesion and democracy in Mexico* [IDOS Discussion Paper 2/2024]. IDOS (German Institute of Development and Sustainability). https://www.idos-research.de/uploads/media/DP_2.2024.pdf
- Brevini, B. (2021). *Is AI good for the planet?* Polity Press.
- Brewster, T. (2024, January 10). Musk's X fired 80% of engineers working on trust and safety, Australian Government says. *Forbes*. <https://www.forbes.com/sites/ssbrewster/2024/01/10/elon-musk-fired-80-per-cent-of-twitter-x-engineers-working-on-trust-and-safety>
- Bria, F. (2018). *Barcelona city council digital plan: Government measure concerning ethical management and accountable data: Barcelona data commons*. Open Digitalisation Programme, Barcelona City Council Office for Technology and Digital Innovation. https://www.barcelona.cat/digitalstandards/en/data-management/0.1/_attachments/barcelona_data_management_0.1.en.pdf
- Bria, F., Timmers, P., & Gernone, F. (2025). *EuroStack – A European alternative for digital sovereignty*. Report commissioned by Bertelsmann Stiftung, Gutersloh. https://www.euro-stack.info/docs/EuroStack_2025.pdf
- Briant, E. L. (2021). *Propaganda machine: The hidden story of Cambridge Analytica and the digital influence industry*. Bloomsbury.
- Briant, E. L. (2024). Hack attacks: How cyber intimidation and conspiracy theories drive the spiral of 'secrecy hacking'. In J. Steel, & J. Petley (Eds.), *The Routledge companion to freedom of expression and censorship* (pp. 285–295). Routledge. <https://doi.org/10.4324/9780429262067>
- Bridgman, A., Merkley, E., Loewen, P. J., Owen, T., Ruths, D., Teichmann, L., & Zhilin, O. (2020). The causes and consequences of COVID-19 misperceptions: Understanding the role of news and social media. *HKS Misinformation Review*, 1(Special Issue), 1–18. <https://misinfoforeview.hks.harvard.edu/article/the-causes-and-consequences-of-covid-19-misperceptions-understanding-the-role-of-news-and-social-media>
- Bright, L. F., Logan, K., & Lim, H. S. (2022). Social media fatigue and privacy: An exploration of antecedents to consumers' concerns regarding the security of their personal information on social media platforms. *Journal of Interactive Advertising*, 22(2), 125–140. <https://doi.org/10.1080/15252019.2022.2051097>
- Brignull, H. (2023). *Deceptive patterns: Exposing the tricks tech companies use to control you*. Testimonium.

- Brin, S., & Page, L. (2012). Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer Networks*, 56(18), 3825–3833. (Original work published 1998) <https://doi.org/10.1016/j.comnet.2012.10.007>
- Broeders, D., Cristiano, F., & Kaminska, M. (2023). In search of digital sovereignty and strategic autonomy: Normative power Europe to the test of its geopolitical ambitions. *JCMS: Journal of Common Market Studies*, 61(5), 1261–1280. <https://doi.org/10.1111/jcms.13462>
- Brogi, E., & Sjøvaag, H. (2023). *Good practices for sustainable new media financing*. Committee of Experts on Increasing Resilience of Media (MSI-RES), Council of Europe. <https://rm.coe.int/msi-res-2022-08-good-practices-for-sustainable-media-financing-for-sub/1680adf466>
- Brogi, E., Borges, D., Carlini, R., Nenadic, I., Bleyer-Simon, K., Kermer, J., Reviglio, U., Trevisan, M., & Verza, S. (2023). *The European media freedom act: Media freedom, freedom of expression and pluralism*. European Parliament's Committee on Civil Liberties, Justice and Home Affairs. [https://www.europarl.europa.eu/RegData/etudes/STUD/2023/747930/IPOL_STU\(2023\)747930_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2023/747930/IPOL_STU(2023)747930_EN.pdf)
- Broniatowski, D. A., Simons, J. R., Gu, J., Jamison, A. M., & Abrams, L. C. (2023). The efficacy of Facebook's vaccine misinformation policies and architecture during the COVID-19 pandemic. *Science Advances*, 9(37), 1–17. <https://doi.org/10.1126/sciadv.adh2132>
- Broomfield, H., & Reutter, L. (2022). In search of the citizen in the datafication of public administration. *Big Data & Society*, 9(1), 1–14. <https://doi.org/10.1177/20539517221089302>
- Broughton Micova, S. B., & Jacques, S. (2020). The functions of data in the competition between audiovisual media and video sharing platforms for advertising. *Journal of Information Policy*, 10, 514–548. <https://doi.org/10.5325/jinfopoli.10.2020.0514>
- Broughton Micova, S., & Schnurr, D. (2024). *Systemic risk in digital services: Benchmarks for Evaluating the management of risks to electoral processes: Report*. CERRE (Centre on Regulation in Europe). https://cerre.eu/wp-content/uploads/2024/05/240523_CERRE_DSA-Systemic-Risk-2_Elections-paper_FINAL.pdf
- Brown, C., & Levin, J. (2020, June 30). *Prioritizing original news reporting on Facebook* [Blog post]. Meta. <https://about.fb.com/news/2020/06/prioritizing-original-news-reporting-on-facebook>
- Brown, I., & Marsden, C. T. (2023). *Regulating code: good governance and better regulation in the information age*. MIT Press.
- Browne, S. (2015). *Dark matters: On the surveillance of blackness*. Duke University Press.
- Bryant, M. (2021, September 26). Is Facebook leading us on a journey to the metaverse? *The Observer*. <https://www.theguardian.com/technology/2021/sep/26/is-facebook-leading-us-on-a-journey-to-the-metaverse>
- Buarque, B., & Zavershinskaia, P. (2022). The far-right politics of 'truth': An exploratory analysis of the 'truths' produced by AfD Kompakt and Patriotic Alternative. In V. A. Bruno (Ed.), *Populism and far-right: Trends in Europe* (pp. 25–48). Educatt.
- Budak, C., Nyhan, B., Rothschild, D. M., Thorson, E., & Watts, D. J. (2024). Misunderstanding the harms of online misinformation. *Nature*, 630(8015), 45–53. <https://doi.org/10.1038/s41586-024-07417-w>
- Bullock, J. B., Chen, Y.-C., Himmelreich, J., Hudson, V. M., A. Korinek, M. M. Young, & B. Zhang (2022). *The Oxford handbook of AI governance*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197579329.001.0001>
- Bunting, M. (2018). *Keeping consumers safe online: Legislating for platform accountability for online content*. Communications Chambers. <https://static1.1.sqspcdn.com/static/f/1321365/27941308/1530714958163/Sky+Platform+Accountability+FINAL+020718+2200.pdf>
- Buolamwini, J. (2023). *Unmasking AI: My mission to protect what is human in a world of machines*. Random House.
- Burgess, J., & Hurcombe, E. (2019). Digital journalism as symptom, response, and agent of change in the platformed media environment. *Digital Journalism*, 7(3), 359–367. <https://doi.org/10.1080/21670811.2018.1556313>
- Burgess, M. (2016, July 28). Google's DeepMind trains AI to cut its energy bills by 40%. *Wired*. <https://www.wired.com/story/google-deepmind-data-centres-efficiency>

- Burkart, N., & Huber, M. F. (2021). A survey on the explainability of supervised machine learning. *Journal of Artificial Intelligence Research*, 70, 245–317. <https://doi.org/10.1613/jair.1.12228>
- Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12. <https://doi.org/10.1177/2053951715622512>
- Business Standard. (2018, August 18). US tech firms to boost lobbying against India’s push to store data locally. www.business-standard.com/article/international/us-tech-firms-to-boost-lobbying-against-india-s-push-to-store-data-locally-118081800348_1.html
- Buturoiu, R., Corbu, N., & Boțan, M. (2023). *Patterns of news consumption in a high-choice media environment: A Romanian perspective*. Springer. <https://doi.org/10.1007/978-3-031-41954-6>
- Caballero, F. S., & Monje, D. I. (2024). Artificial intelligence and technological accelerationism: A critique of cybernetic ideology. In J. Pedro-Caranana, R. Gómez, T. F. Corrigan, & F. S. Caballero (Eds.), *Political economy of media and communication: Methodological approaches* (pp. 328–365). Routledge. <https://doi.org/10.4324/9781003385516>
- CAIDP (Center for AI and Digital Policy). (2025). *AI and democratic values index 2025: A comprehensive review of AI policies and practices worldwide*. <https://www.caidp.org/reports/aidv-2025/>
- Calderaro, A., & Blumfelde, S. (2022). Artificial intelligence and EU security: The false promise of digital sovereignty. *European Security*, 31(3), 415–434. <https://doi.org/10.1080/09662839.2022.2101885>
- Calhoun, C., Gaonkar, D. P., & Taylor, C. (2022). *Degenerations of democracy*. Harvard University Press.
- Calice, M. N., Bao, L., Freiling, I., Howell, E., Xenos, M. A., Yang, S., Brossard, D., Newman, T. P., & Scheufele, D. A. (2023). Polarized platforms? How partisanship shapes perceptions of “algorithmic news bias”. *New Media & Society*, 25(11), 2833–2854. <https://doi.org/10.1177/14614448211034159>
- Calo, R., & Citron, D. K. (2021). The automated administrative state: A crisis of legitimacy. *Emory Law Journal*, 70(4), 799–844. <https://scholarlycommons.law.emory.edu/elj/vol70/iss4/1>
- Camara, M. S., Banu, H., & Abeck, J. C. (2023). The African youth and social media at the crossroads of information, misinformation, and disinformation. In K. Langmia (Ed.), *Black communication in the age of disinformation: DeepFakes and synthetic media* (pp. 15–50). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-27696-5_2
- Cammaerts, B. (2018). *The circulation of anti-austerity protest*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-70123-3>
- Cammaerts, B. (2024). Defending democracy against populist neo-Fascist attacks: The role and problems of public sphere theory. *Javnost – The Public*, 31(1), 26–45. <https://doi.org/10.1080/13183222.2024.2310983>
- Cammaerts, B., & Mansell, R. (2020). Digital platform policy and regulation: Toward a radical democratic turn. *International Journal of Communication*, 14, 135–154. <https://ijoc.org/index.php/ijoc/article/view/11182/2901>
- Campbell-Stephens, R. M. (2021). *Educational leadership and the global majority: Decolonising narratives*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-88282-2>
- Cano-Orón, L., Calvo, D., García, G. L., & Baviera, T. (2021). Disinformation in Facebook ads in the 2019 Spanish general election campaigns. *Media and Communication*, 9(1), 217–228. <https://doi.org/10.17645/mac.v9i1.3335>
- Caplan, R. (2023). Networked platform governance: The construction of the democratic platform. *International Journal of Communication*, 17, 3541–3472. <https://ijoc.org/index.php/ijoc/article/view/20035>
- Carmi, E., & Yates, S. (2023). Data citizenship: Data literacies to challenge power imbalance between society and “big tech”. *International Journal of Communication*, 17, 3619–3637. <https://ijoc.org/index.php/ijoc/article/view/18823>
- Carroll, S. R., Rodriguez-Lonebear, D., & Martinez, A. (2019). Indigenous data governance: Strategies from the United States’ Native nations. *Data Science Journal*, 18(31), 1–15. <https://doi.org/doi:10.5334/dsj-2019-031>

- Carson, A., & Gibbons, A. (2023). The big chill? How journalists and sources perceive and respond to fake news laws in Indonesia and Singapore. *Journalism Studies*, 24(14), 1819–1838. <https://doi.org/10.1080/1461670X.2023.2192299>
- Carson, A., & Wright, S. (2022). Fake news and democracy: Definitions, impact and response. *Australian Journal of Political Science*, 57(3), 221–230. <https://doi.org/10.1080/10361146.2022.2122778>
- Casati, R. (2013). *Contro Il Colonialismo Digitale [Against digital colonialism]*. Laterza.
- Casciani, D., & BBC Verify. (2024, August 2). *Violent Southport protests reveal far-right organising tactics*. BBC News. <https://www.bbc.com/news/articles/cl4y0453nv5o>
- Casero-Ripollés, A., Tuñón, J., & Bouza-García, L. (2023). The European approach to online disinformation: Geopolitical and regulatory dissonance. *Humanities and Social Sciences Communications*, 10(1), 1–10. <https://doi.org/10.1057/s41599-023-02179-8>
- Castro, L., Strömbäck, J., Esser, F., Van Aelst, P., de Vreese, C., Aalberg, T., Cardenal, A. S., Corbu, N., Hopmann, D. N., Koc-Michalska, K., Matthes, J., Schemer, C., Sheaffer, T., Splendore, S., Stanyer, J., Stępińska, A., Štětka, V., & Theocharis, Y. (2022). Navigating high-choice European political information environments: A comparative analysis of news user profiles and political knowledge. *The International Journal of Press/Politics*, 27(4), 827–859. <https://doi.org/10.1177/19401612211012572>
- Catalán-Matamoros, D. (Ed.) (2023). *Disinformation and fact-checking in contemporary society*. ESIC.
- Caton, S., & Haas, C. (2020). Fairness in machine learning: A survey. *ArXiv*. <https://doi.org/10.48550/arXiv.2010.04053>
- Caulfield, M., Bayar, M. C., & Aske, A. B. (2023). *The ‘new elites’ of X: Identifying the most influential accounts engaged in Hamas/Israel discourse*. Center for an Informed Public, University of Washington. <https://www.cip.uw.edu/2023/10/20/new-elites-twitter-x-most-influential-accounts-hamas-israel/>
- Center for Countering Digital Hate. (2021). *The toxic ten: How ten fringe publishers fuel 69% of digital climate change denial*. <https://counterhate.com/research/the-toxic-ten>
- CERTH (Center for Research and Technology Hellas). (2021). *Online disinformation in Europe: A study on ethical standards, involved political interests and policy recommendations*. The Left in the European Parliament. <https://left.eu/issues/study-online-disinformation-in-europe>
- Chadha, K., & Bhat, P. (2022). Alternative news media and critique of mainstream journalism in India: The case of OpIndia. *Digital Journalism*, 10(8), 1283–1301. <https://doi.org/10.1080/21670811.2022.2118143>
- Chadwick, A. (2017). *The hybrid media system: Politics and power* (2nd ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780190696726.001.0001>
- Chaguaceda, A., Cilano Pelaez, J., & Puerta, M. I. (2023). Illiberal narratives in Latin America: Russian and allied media as vehicles of autocratic cooperation. *The Journal of Illiberalism Studies*, 3(2), 111–123. <https://doi.org/10.53483/XCMW3558>
- Chair, C. (2024). Gendered digital inequalities: How do we ensure gender transformative law and practice in the age of artificial intelligence in Africa? In L. A. Abdulrauf, & H. Dube (Eds.), *Data privacy law in Africa: Emerging perspectives* (pp. 207–238). Pretoria University Law Press.
- Chairpersons Organizations. (2023). *Statement of the networks of media and online safety regulators on the occasion of the UNESCO global conference internet for trust*. ACRAN (African Communication Regulation Authorities Network); ERGA (European Regulators Group for Audiovisual Media Services); GOSRN (Global Online Safety Regulators Network); MNRA (Mediterranean Network of Regulatory Authorities); PRAI (Platform of Ibero-American Audiovisual Regulators); REFRAM (Francophone Network of Media Regulators). <https://www.rirm.org/wp-content/uploads/2017/06/Statement-UNESCO-Global-Conference-en.pdf>
- Chambers, R. (1997). *Whose reality counts? Putting the first last*. Practical Action Publishing.
- Chambers, S. (2023). Deliberative democracy and the digital public sphere: Asymmetrical fragmentation as a political not a technological problem. *Constellations*, 30(1), 61–68. <https://doi.org/10.1111/1467-8675.12662>
- Chasi, C., & Rodney-Gumede, Y. (2022). Decolonial journalism: New notes on ubuntu and the public interest. *Journalism Studies*, 23(13), 1625–1637. <https://doi.org/10.1080/1461670X.2022.2083006>

- Chavalarias, D. (2024). *Minuit moins dix à l'horloge de Poutine: Analyse de réseaux des ingérences étrangères dans les élections législatives de 2024* [Ten to midnight on Putin's clock: A network analysis of foreign interference in the 2024 parliamentary elections]. Hal Open Science. <https://hal.science/hal-04629585v3/document>
- Chen, J. Q., & Benusa, A. (2017). HIPAA security compliance challenges: The case for small healthcare providers. *International Journal of Healthcare Management*, 10(2), 135–146. <https://doi.org/10.1080/20479700.2016.1270875>
- Chen, Y., & Gao, Y. (2022). Comparative analysis of digital trade development strategies and governance approaches. *Journal of Digital Economy*, 1(3), 227–238. <https://doi.org/10.1016/j.jdec.2023.02.001>
- Cheng, D. (2016). *Cyber dragon: Inside China's information warfare and cyber operations*. Bloomsbury. <https://doi.org/10.5040/9798400636431>
- Chi, N., Lurie, E., & Mulligan, D. K. (2021). Reconfiguring diversity and inclusion for AI ethics. *Proceedings of the AIES '21: AAAI/ACM Conference on AI, Ethics, and Society* (pp. 447–457). <https://doi.org/10.1145/3461702.3462622>
- Chibwe, A. (2020). Social media and elections in Zimbabwe: Twitter war between pro-ZANU-PF and pro-MDC-A netizens. *South African Journal of Communication Theory and Research*, 46(4), 7–30. <https://doi.org/10.1080/02500167.2020.1723663>
- Chin, Y. C. (2024). Right to data access in the digital era: The case of China. In M. Aslama Horowitz, H. Niemeinen, K. Lehtisaari, & A. D'Arma (Eds.), *Epistemic rights in the era of digital disruption* (pp. 187–201). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-45976-4_13
- China Ministry of Education. (2019). AI innovation action plan for institutions of higher education (trans. Center for security and emerging technology), November 14. <https://cset.georgetown.edu/publication/ai-innovation-action-plan-for-institutions-of-higher-education/>
- Chiumbu, S., & Munoriyarwa, A. (2023). Exploring data journalism practices in Africa: Data politics, media ecosystems and newsroom infrastructures. *Media, Culture & Society*, 45(4), 841–858. <https://doi.org/10.1177/01634437231155341>
- Cho, H., Cannon, J., Lopez, R., & Li, W. (2024). Social media literacy: A conceptual framework. *New Media & Society*, 26(2), 941–960. <https://doi.org/10.1177/14614448211068530>
- Choi, J., Kim, K. K., & Kim, Y. (2020). Emotional analysis with news using text mining for framing theory. In R. Lee (Ed.), *Computer and information science* (pp. 95–108). Springer. https://doi.org/10.1007/978-3-030-25213-7_7
- Chonka, P., Diepeveen, S., & Haile, Y. (2023). Algorithmic power and African indigenous languages: Search engine autocomplete and the global multilingual Internet. *Media, Culture & Society*, 45(2), 246–265. <https://doi.org/10.1177/01634437221104705>
- Chouliaraki, L., & Georgiou, M. (2022). *The digital border: Migration, technology, power*. NYU Press.
- Chow, A. R. (2025, January 7). Why Meta's fact-checking change could lead to more misinformation on Facebook and Instagram. *TIME*. <https://time.com/7205332/meta-fact-checking-community-notes/>
- Christodoulou, E., & Iordanou, K. (2021). Democracy under attack: Challenges of addressing ethical issues of AI and big data for more democratic digital media and societies. *Frontiers in Political Science*, 3, 1–17. <https://doi.org/10.3389/fpos.2021.682945>
- Christofoletti, R., & Becker, D. (2023). Pesquisas sobre confiança na mídia e o fortalecimento da ideia de crise permanente de credibilidade [Research on trust in the media and the strengthening of the idea of a permanent crisis of credibility]. *Chasqui: Revista Latinoamericana de Comunicación*, 154, 53–67. <https://dialnet.unirioja.es/servlet/articulo?codigo=9274150>
- Christopher, N. (2021, October 25). Amazon's "delivery premier league" gamifies gig work in India. Rest of World Reporting Global Tech Stories. <https://restofworld.org/2021/amazons-delivery-premier-league-gamifies-gig-work-in-india>
- Chung, M., & Wihbey, J. (2024). Social media regulation, third-person effect, and public views: A comparative study of the United States, the United Kingdom, South Korea, and Mexico. *New Media & Society*, 26(8), 4534–4553. <https://doi.org/10.1177/14614448221122996>

- Church, P., & Pehlivan, C. N. (2023). The digital services act (DSA): A new era for online harms and intermediary liability. *Global Privacy Law Review*, 4(1), 53–59. <https://doi.org/10.54648/gplr2023005>
- Cieslik, K., & Margócsy, D. (2022). Datafication, power and control in development: A historical perspective on the perils and longevity of data. *Progress in Development Studies*, 22(4), 352–373. <https://doi.org/10.1177/14649934221076580>
- Cipers, S., Meyer, T., & Lefevere, J. (2023). Government responses to online disinformation unpacked. *Internet Policy Review*, 12(4). <https://doi.org/10.14763/2023.4.1736>
- Cities Coalition for Digital Rights. (n.d.). *About us*. Retrieved June 30, 2025, from <https://citiesfordigitalrights.org/thecoalition>
- Claesson, A. (2023). Twitter: A necessary evil? Journalistic responses to Elon Musk and the denormalization of social media. *Journalism*, 25(12), 2604–2621. <https://doi.org/10.1177/14648849231221616>
- CNTI (Center for News, Technology & Innovation). (2024, September 3). *Most “fake news” legislation risks doing more harm than good amid a record number of elections in 2024*. <https://innovating.news/article/most-fake-news-legislation-risks-doing-more-harm-than-good-amid-a-record-number-of-elections-in-2024>
- Cohen, J. E. (2019). *Between truth and power: The legal constructions of informational capitalism*. Oxford University Press. <https://doi.org/10.1093/oso/9780190246693.001.0001>
- Cohen, J. E. (2023). Infrastructuring the digital public sphere. *Yale Journal of Law and Technology*, 25(1), 1–40. <https://yolt.org/infrastructuring-digital-public-sphere>
- Cohn, A., Stoop, J., & Rahman, H. A. (2022). *Disinformation for hire: Examining the production of false COVID-19 information* [Tinbergen Institute Discussion Paper]. <https://papers.tinbergen.nl/22086.pdf>
- Colclough, C. (2022). Reshaping the digitization of public services. *New England Journal of Public Policy*, 34(1), 1–14. <https://scholarworks.umb.edu/nejpp/vol34/iss1/9>
- Cole, M. D. (2024). AI regulation and governance on a global scale: An overview of international, regional and national instruments. *Journal of AI Law and Regulation*, 1(1), 126–142. <https://doi.org/10.21552/aire/2024/1/16>
- Collao, K. (2024). *Ok computer? Public attitudes to the uses of generative AI in news*. Craft; Reuters Institute for the Study of Journalism, Oxford University. <https://reutersinstitute.politics.ox.ac.uk/news/ok-computer-understanding-public-attitudes-towards-uses-generative-ai-news>
- Conroy-Krutz, J., & Koné, J. (2022). *Confronting threats to civic spaces: Final report*. Afrobarometer. <https://www.afrobarometer.org/wp-content/uploads/2023/05/NED-Confronting-Threats-Report-Afrobarometer-Final-27june22.pdf>
- Cooke, D. (2023). *Synthetic media and election integrity: Defending our democracies*. Alan Turing Institute, Centre for Emerging Technology and Security (CETaS). <https://cetas.turing.ac.uk/publications/synthetic-media-and-election-integrity-defending-our-democracies>
- Cooke, P. (2020). Silicon Valley imperialists create new model villages as smart cities in their own image. *Journal of Open Innovation: Technology, Market, and Complexity*, 6(24), 1–18. <https://doi.org/10.3390/joitmc6020024>
- Coombs, W. T., & Holladay, S. J. (Eds.). (2022). *The handbook of crisis communication* (2nd ed.). John Wiley & Sons. <https://doi.org/10.1002/9781119678953>
- Coos, A. (2022, April 5). *Data protection in Japan: All you need to know about APPI* [Blog post]. Endpoint Protector. www.endpointprotector.com/blog/data-protection-in-japan-appi
- Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic decision making and the cost of fairness. *ArXiv*. <https://doi.org/10.48550/ArXiv.1701.08230>
- Corbu, N., Bârgăoanu, A., Durach, F., & Udrea, G. (2021). Fake news going viral: The mediating effect of negative emotions. *Media Literacy and Academic Research*, 4(2), 58–87. https://www.mlar.sk/wp-content/uploads/2021/12/4_Corbu.pdf
- Cornils, M. (2020). *Designing platform governance: A normative perspective on needs, strategies, and tools to regulate intermediaries*. AlgorithmWatch; European Policy Centre; Mainzer Medieninstitut and Civitates. <https://algorithmwatch.org/de/wp-content/uploads/2020/05/Governing-Platforms-legal-study-Cornils-May-2020-AlgorithmWatch.pdf>

- Corporate Europe Observatory. (2024a, March 12). *Don't let corporate lobbying further water down the AI act, lobby watchdogs warn MEPs*. <https://corporateeurope.org/en/2024/03/dont-let-corporate-lobbying-further-water-down-ai-act-lobby-watchdogs-warn-meps>
- Corporate Europe Observatory. (2024b, March 11). *Trojan horses: How European startups teamed up with big tech to gut the AI act*. <https://www.corporateeurope.org/en/2024/03/trojan-horses-how-european-startups-teamed-big-tech-gut-ai-act>
- Corrales Compagnucci, M., Wilson, M. L., Fenwick, M., Forgó, N., & Bärnighausen, T. (Eds.). (2022). *AI in eHealth: Human autonomy, data governance and privacy in healthcare*. Cambridge University Press.
- Couldry, N., & Curran, J. (Eds.). (2003). *Contesting media power: Alternative media in a networked world*. Rowman & Littlefield.
- Couldry, N., & Hepp, A. (2016). *The mediated construction of reality: Society, culture, mediatization*. Polity.
- Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, 20(4), 336–349. <https://doi.org/10.1177/1527476418796632>
- Council of Europe. (2018). *Recommendation CM/Rec(2018)1 of the committee of ministers to member states on media pluralism and transparency of media ownership*. <https://search.coe.int/cm?i=0900001680790e13>
- Council of Europe. (2022). *Recommendation CM/Rec(2022)12 of the committee of ministers to member states on electoral communication and media coverage of election campaigns*. <https://search.coe.int/cm?i=0900001680a6172e>
- Council of Europe. (2023). *Guidelines for the responsible implementation of artificial intelligence (AI) in journalism*. <https://www.coe.int/en/web/freedom-expression/-/guidelines-on-the-responsible-implementation-of-artificial-intelligence-ai-systems-in-journalism>
- Council of Europe. (2024, May 17). *Framework Convention on artificial intelligence, human rights, democracy and the rule of law*. <https://search.coe.int/cm?i=0900001680afb122>
- Cows, J., Tsamados, A., Taddeo, M., & Floridi, L. (2023). The AI gambit: Leveraging artificial intelligence to combat climate change – Opportunities, challenges, and recommendations. *AI & Society*, 38(1), 283–307. <https://doi.org/10.1007/s00146-021-01294-x>
- Craig, J., Simmons, C., & Bhatnagar, R. (2023, January 13). *How January 6 inspired election disinformation around the world*. ISD Digital Dispatches. https://www.isdglobal.org/digital_dispatches/how-january-6-inspired-election-disinformation-around-the-world
- Cramer, B. W. (2020). From liability to accountability: The ethics of citing section 230 to avoid the obligations of running a social media platform. *Journal of Information Policy*, 10, 123–150. <https://doi.org/10.5325/jinfopoli.10.2020.0123>
- Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Crawford, K., Gray, M. L., & Miltner, K. (2014). Critiquing big data: Politics, ethics, epistemology. *International Journal of Communication*, 8, 1663–1672. <https://ijoc.org/index.php/ijoc/article/view/2167>
- Crémer, J., de Montjoye, Y.-A., & Schweitzer, H. (2019). *Competition policy for the digital era: Final report*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2763/407537>
- Crosset, V., & Dupont, B. (2022). Cognitive assemblages: The entangled nature of algorithmic content moderation. *Big Data & Society*, 9(2), 1–13. <https://doi.org/10.1177/20539517221143361>
- Crump, C. (2016). Surveillance policy making by procurement. *Washington Law Review*, 91(4), 1595–1662. <https://doi.org/10.2139/ssrn.2737006>
- Culloty, E. (2021). Désinformation sur la migration: UN problème de longue date au dimensions technologiques nouvelles [Migration disinformation: A long-standing problem with new technological dimensions]. In M. McAuliffe, & A. Triandafyllidou (Eds.), *Rapport état de la migration dans le monde [World migration report]* (pp. 229–245). Organisation internationale pour les migrations (OIM).

- Cunliffe-Jones, P. (2021). How government responses to misinformation in Africa restrict freedom of expression and do little to tackle the problem. *African Journalism Studies*, 42(2), 121–125. <https://doi.org/10.1080/23743670.2021.1972532>
- Curran, J., & Hesmondhalgh, D. (Eds.). (2019). *Media and society* (6th ed.). Bloomsbury.
- Curran, J., & Park, M.-J. (Eds.). (1999). *De-westernizing media studies*. Routledge. <https://doi.org/10.4324/9780203981764>
- Curtis, J. S. (2021). Springing the ‘Tacitus trap’: Countering Chinese state-sponsored disinformation. *Small Wars & Insurgencies*, 32(2), 229–265. <https://doi.org/10.1080/09592318.2021.1870429>
- Curzi de Mendonça, Y. (2024, September 9). Elon Musk’s feud with Brazilian judge is much more than a personal spat – it’s about national sovereignty, freedom of speech and the rule of law. *The Conversation*. <https://theconversation.com/elon-musks-feud-with-brazilian-judge-is-much-more-than-a-personal-spat-its-about-national-sovereignty-freedom-of-speech-and-the-rule-of-law-238264>
- Cushion, S. (2021). Are public service media distinctive from the market? Interpreting the political information environments of BBC and commercial news in the United Kingdom. *European Journal of Communication*, 37(1), 3–20. <https://doi.org/10.1177/02673231211012149>
- Custers, B., & Fosch-Villaronga, E. (Eds.). (2022). *Law and artificial intelligence: Regulating AI and applying AI in legal practice*. T.M.C. Asser Press.
- Cyphers, B. (2022, August 31). *Inside fog data science, the secretive company selling mass surveillance to local police*. Electronic Frontier Foundation. <https://www.eff.org/deeplinks/2022/08/inside-fog-data-science-secretive-company-selling-mass-surveillance-local-police>
- D’Haenens, L., Lo, W. H., & Moore, M. (2022). Innovation in journalism: How technology affects the news media, publication formats, and the journalist profession. In J. Trappel, & T. Tomaz (Eds.), *Success and failure in news media performance: Comparative analysis in the media for democracy monitor 2021* (pp. 337–345). Nordicom, University of Gothenburg. <https://doi.org/10.48335/9789188855589-16>
- da Silva, C. R. S., Teixeira, T. M. C., & Pinto, V. B. (2019). Research methodology in information literacy: A systematic review. *RDBCI: Revista Digital de Biblioteconomia e Ciência Da Informação*, 17, 1–17. <https://doi.org/10.20396/rdbci.v17i0.8653728/e019014>
- Dadakhonov, A. O. (2024). Analysis of media and information literacy definitions: A qualitative approach. *Studies in Media and Communications*, 12(2), 116–129. <https://doi.org/10.11114/smc.v12i2.6659>
- Dahlberg, L. (2014). The Habermasian public sphere and exclusion: An engagement with poststructuralist-influenced critics. *Communication Theory*, 24(1), 21–41. <https://doi.org/10.1111/comt.12010>
- DAIR (Distributed Artificial Intelligence Research) Institute. (n.d.). *Our research is intended to benefit communities*. Retrieved July 4, 2025, from <https://www.dair-institute.org/research-philosophy>
- Dalle, J.-M., & David, P. A. (2007). Simulating code growth in Libre (open source) mode. In E. Brousseau, & N. Curien (Eds.), *Internet and digital economics: Principles, methods and applications* (pp. 391–422). Cambridge University Press. <https://doi.org/10.1017/CBO9780511493201>
- Daniels, J. (2018). The algorithmic rise of the “alt-right”. *Contexts*, 17(1), 60–65. <https://doi.org/10.1177/1536504218766547>
- Daniels, D. (2025, March 18). *Why USAID cuts highlight the role of public media*. Public Media Alliance. <https://www.publicmediaalliance.org/why-usaid-cuts-highlight-the-role-of-public-media/>
- Das, A., & Schroeder, R. (2021). Online disinformation in the run-up to the Indian 2019 election. *Information, Communication & Society*, 24(12), 1762–1778. <https://doi.org/10.1080/1369118X.2020.1736123>
- Davies, M., & Vipra, J. (2025). *Computing commons: Designing public compute for people and society*. Ada Lovelace Institute. <https://www.adalovelaceinstitute.org/report/computing-commons/>
- Davis, A., Fenton, N., Freedman, D., & Khiabany, G. (2020). *Media, democracy and social change: Re-imagining political communications*. Sage. <https://doi.org/10.4135/9781529730166>

- De Abreu, B. S. (Ed.). (2022). *Media literacy, equity, and justice*. Routledge. <https://doi.org/10.4324/9781003175599>
- De Blasio, E., & Selva, D. (2021). Who is responsible for disinformation? European approaches to social platforms' accountability in the post-truth era. *American Behavioral Scientist*, 65(6), 825–846. <https://doi.org/10.1177/0002764221989784>
- De Bruyne, J., & Vanleenhove, C. (Eds.). (2021). *Artificial intelligence and the law*. Intersentia. Deceptive Patterns. (n.d.). *Deceptive patterns* [Web page]. Retrieved September 2, 2025 from <https://www.deceptive.design/>
- De Gregorio, G. (2023). The normative power of artificial intelligence. *Indiana Journal of Global Legal Studies*, 30(2), 55–80. <https://ssrn.com/abstract=4436287>
- De Gregorio, G., & Dunn, P. (2023). Artificial intelligence and freedom of expression. In A. Quintavalla, & J. Temperman (Eds.), *Artificial intelligence and human rights* (pp. 76–90). Oxford University Press. <https://doi.org/10.1093/law/9780192882486.001.0001>
- De Gregorio, G., & Stremlau, N. (2021). Platform governance at the periphery: Moderation, shutdowns and intervention. In J. Bayer, B. Holznagel, P. Korpisaari, & L Woods (Eds.), *Perspectives on platform regulation: Concepts and models of social media governance across the globe* (pp. 433–450). Nomos. <https://doi.org/10.5771/9783748929789>
- De Gregorio, G., & Stremlau, N. (2023). Inequalities and content moderation. *Global Policy*, 14(5), 870–879. <https://doi.org/10.1111/1758-5899.13243>
- de-Lima Santos, M.-F. (2023). The entanglements between data journalism, collaboration and business models: A systematic literature review. *Digital Journalism*, 12(2), 256–281. <https://doi.org/10.1080/21670811.2023.2247449>
- de Lima Santos, M.-F., Yeung, W. N., & Dodds, T. (2024). Guiding the way: A comprehensive examination of AI guidelines in global media. *AI & SOCIETY*, 40, 2585–2603. <https://doi.org/10.1007/s00146-024-01973-5>
- De Paor, S., & Heravi, B. (2020). Information literacy and fake news: How the field of librarianship can help combat the epidemic of fake news. *The Journal of Academic Librarianship*, 46(5), 1–8. <https://doi.org/10.1016/j.acalib.2020.102218>
- de Sousa Santos, B. (2018). *The end of the cognitive empire: The coming of age of epistemologies of the south*. Duke University Press.
- de Vivo, I. (2023). The “neo-intermediation” of large on-line platforms: Perspectives of analysis of the ‘state of health’ of the digital information ecosystem. *Communications*, 48(3), 420–439. <https://doi.org/10.1515/commun-2022-0102>
- Dellinger, J., & Pell, S. (2024). Bodies of evidence: The criminalization of abortion and surveillance of women in a post-DOBBs world. *Duke Journal of Constitutional Law & Public Policy*, 19, 1–99. <http://dx.doi.org/10.2139/ssrn.4599445>
- Demeter, M. (2019). The winner takes it all: International inequality in communication and media studies today. *Journalism & Mass Communication Quarterly*, 96(1), 37–59. <https://doi.org/10.1177/1077699018792270>
- Dencik, L. (2025). ‘Rescuing’ data justice? Mobilising the collective in responses to datafication. *Information, Communication & Society*, 28(6), 1023–1038. <https://doi.org/10.1080/1369118X.2025.2465874>
- Dencik, L., & Leistert, O. (Eds.). (2015). *Critical perspectives on social media and protest: Between control and emancipation*. Rowman & Littlefield.
- Dencik, L., Hintz, A., & Cable, J. (2016). Towards data justice? The ambiguity of anti-surveillance resistance in political activism. *Big Data & Society*, 3(2), 1–12. <https://doi.org/10.1177/2053951716679678>
- Dergacheva, D., Kuznetsova, V., Scharlach, R., & Katzenbach, C. (2023). *One day in content moderation: Analyzing 24 h of social media platforms' content decisions through the DSA transparency database*. University of Bremen. <https://doi.org/10.26092/elib/2707>
- Deuze, M. (2014). *Media life*. Polity Press.
- Deuze, M., & Beckett, C. (2022). Imagination, algorithms and news: Developing AI literacy for journalism. *Digital Journalism*, 10(10), 1913–1918. <https://doi.org/10.1080/21670811.2022.2119152>
- Devenney, M. (2009). The limits of communicative rationality and deliberative democracy. *Journal of Power*, 2(1), 137–154. <https://doi.org/10.1080/17540290902760915>

- Dias Oliva, T. (2020). Content moderation technologies: Applying human rights standards to protect freedom of Expression. *Human Rights Law Review*, 20(4), 607–640. <https://doi.org/10.1093/hrlr/ngaa032>
- Diaz Ruiz, C. (2023). Disinformation on digital media platforms: A market-shaping approach. *New Media & Society*, 27(4), 2188–2211. <https://doi.org/10.1177/14614448231207644>
- Díaz, Á., & Hecht-Felella, L. (2021). *Double standards in social media content moderation*. Brennan Center for Justice, New York University School of Law. https://www.brennancenter.org/sites/default/files/2021-08/Double_Standards_Content_Moderation.pdf
- Díaz, A., & Levinson-Waldman, R. (2020). *Automatic license plate readers: Legal status and policy recommendations for law enforcement use*. Brennan Center for Justice, New York University School of Law. <https://www.brennancenter.org/our-work/research-reports/automatic-license-plate-readers-legal-status-and-policy-recommendations>
- Dibenedetto, C. (2021). *Faced with limited resources, Indigenous communities built their own internet: Here's how*. Indigenous Peoples Major Group for Sustainable Development. <https://indigenouspeoples-sdg.org/index.php/english/ttt/1549-faced-with-limited-resources-indigenous-communities-built-their-own-internet-here-s-how>
- Dierickx, L., Lindén, C.-G., & Opdahl, A. L. (2023c). The information disorder level (IDL) index: A human-based metric to assess the factuality of machine-generated content. In D. Ceolin, T. Caselli, & M. Tulin (Eds.), *Disinformation in open online media* (pp. 60–71). Springer.
- Dierickx, L., Lindén, C.-G., & Opdahl, A. L. (2023a). Automated fact-checking to support professional practices: Systematic literature review and meta-analysis. *International Journal of Communication*, 17, 5170–5190. <https://ijoc.org/index.php/ijoc/article/view/21071>
- Dierickx, L., Linden, C.-G., Opdahl, A., Khan, S. A., & Rojas, D. (2023b). AI in the newsroom: A data quality assessment framework for employing machine learning in journalistic workflows. *Proceedings of the 5th International Conference on Advanced Research Methods and Analytics*. <http://hdl.handle.net/10251/201692>
- DiGi (Digital Industry Group). (2022). *Australian code of practice on disinformation and misinformation*. <https://digi.org.au/disinformation-code>
- DiResta, R. (2024, June 15). My encounter with the fantasy-industrial complex. *The Atlantic*. <https://www.theatlantic.com/ideas/archive/2024/06/cia-renee-censorship-conspiracy-twitter/678688>
- Dittrich, P.-J. (2019). *Tackling the spread of disinformation: Why a co-regulatory approach is the right way forward for the EU*. Bertelsmann Stiftung Policy Paper 12, Hertie School, Jacques Delors Centre. <https://aei.pitt.edu/102500/>
- Dolata, M., Feuerriegel, S., & Schwabe, G. (2022). A sociotechnical view of algorithmic fairness. *Information Systems Journal*, 32(4), 754–818. <https://doi.org/10.1111/isj.12370>
- Domingos Cordeiro, V., & Cozman, F. (2024). Artificial intelligence and everyday knowledge. In H. S. Dunn, M. Ragnedda, M. L. Riuu, & L. Robinson (Eds.), *The Palgrave handbook of everyday digital life* (pp. 23–35). Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-30438-5>
- Dommett, K. (2023). The inter-institutional impact of digital platform companies on democracy: A case study of the UK media's digital campaigning coverage. *New Media & Society*, 25(10), 2763–2780. <https://doi.org/10.1177/14614448211028546>
- Douai, A. (2019). Global and Arab media in the post-truth era: Globalization, authoritarianism and fake news. In *IEMed: Mediterranean Yearbook 2019* (pp. 124–132). IEMed – European Institute of the Mediterranean. <https://www.iemed.org/wp-content/uploads/2021/01/Global-and-Arab-Media-in-the-Post-truth.pdf>
- Douek, E. (2024). The Meta oversight board and the empty promise of legitimacy. *Harvard Journal of Law & Technology*, 37(2), 373–445. <http://dx.doi.org/10.2139/ssrn.4565180>
- Dowling, M.-E. (2022). Cyber information operations: Cambridge Analytica's challenge to democratic legitimacy. *Journal of Cyber Policy*, 7(2), 230–248. <https://doi.org/10.1080/23738871.2022.2081089>
- DPGA (Digital Public Goods Alliance). (2023). *Collaborative momentum: The 2023 state of the digital public goods ecosystem report*. <https://digitalpublicgoods.net/blog/collaborative-momentum-the-2023-state-of-the-digital-public-goods-ecosystem-report/>
- Dubois, E., & Blank, G. (2018). The echo chamber is overstated: The moderating effect of political interest and diverse media. *Information, Communication & Society*, 21(5), 729–745. <https://doi.org/10.1080/1369118X.2018.1428656>

- Duffy, A., & Ang, P. H. (2019). Digital journalism: Defined, refined, or re-defined. *Digital Journalism*, 7(3), 378–385. <https://doi.org/10.1080/21670811.2019.1568899>
- Dugan, L., Ippolito, D., Kirubarajan, A., Shi, S., & Callison-Burch, C. (2023). Real or fake text? Investigating human ability to detect boundaries between human-written and machine-generated text. *ArXiv*. <https://doi.org/10.48550/arXiv.2212.12672>
- Duke Reporters' Lab. (2025). *Fact-checking sites around the world*. Sanford School of Public Policy, Duke University. Retrieved August 28, 2025, from <https://reporterslab.org/fact-checking/>
- Duncan, J. (2023). Data protection beyond data rights: Governing data production through collective intermediaries. *Internet Policy Review*, 12(3), 1–22. <https://doi.org/10.14763/2023.3.1722>
- Dunn, H. S., Ragnedda, M., Ruiui, M. L., & Robinson, L. (Eds.). (2024). *The Palgrave handbook of everyday digital life*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-30438-5>
- Durach, F., Bărgăoanu, A., & Nastasiu, C. (2020). Tackling disinformation: EU regulation of the digital space. *Romanian Journal of European Affairs*, 20(1), 5–20. <https://papers.ssrn.com/abstract=3650780>
- Durach, F., Vladu, L., Călin, I., & Enache, M. (2024). Public policies to fight disinformation actively through media and information literacy. In D. Frau-Meigs, & N. Corbu (Eds.), *Disinformation debunked: Building resilience through media and information literacy* (pp. 257–293). Routledge. <https://doi.org/10.4324/9781003387404>
- Dutta, M. J., & Pal, M. (2020). Theorizing from the Global South: Dismantling, resisting, and transforming communication theory. *Communication Theory*, 30(4), 349–369. <https://doi.org/10.1093/ct/qtaa010>
- Dutta, S., & Lanvin, B. (2023). *Network readiness index 2023: Trust in a network society: A crisis of the digital age?* Portulans Institute, University of Oxford. <https://networkreadinessindex.org/>
- Dutton, W. H. (2023). *The fifth estate: The power shift of the digital age*. Oxford University Press. <https://doi.org/10.1093/oso/9780190688363.001.0001>
- Dyson, I. (2023, December 15). New York City must strengthen police transparency law. Brennan Center for Justice, New York University School of Law. <https://www.brennancenter.org/our-work/analysis-opinion/new-york-city-must-strengthen-police-transparency-law>
- Eady, G., Nagler, J., Guess, A., Zilinsky, J., & Tucker, J. A. (2019). How many people live in political bubbles on social media? Evidence from linked survey and Twitter data. *SAGE Open*, 9(1), 1–21. <https://doi.org/10.1177/2158244019832705>
- Eady, G., Paskhalis, T., Zilinsky, J., Bonneau, R., Nagler, J., & Tucker, J. A. (2023). Exposure to the Russian internet research agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior. *Nature Communications*, 14(1), 1–11. <https://doi.org/10.1038/s41467-022-35576-9>
- Eady, G., Bonneau, R., Tucker, J. A., & Nagler, J. (2024). News sharing on social media: Mapping the ideology of news media, politicians, and the mass public. *Political Analysis*, 33(2), 73–90. <https://doi.org/10.1017/pan.2024.19>
- Eastwood, B. (2024, May 18). Study: Industry now dominates AI research. MIT Management Sloan School. <https://mitsloan.mit.edu/ideas-made-to-matter/study-industry-now-dominates-ai-research>
- EC (European Commission). (2000). *Charter of fundamental rights on the European Union*. C 364/1. https://www.europarl.europa.eu/charter/pdf/text_en.pdf
- EC (European Commission). (2012). *Consolidated version of the treaty on the functioning of the European Union (TFEU)*. C326/47. <https://www.legislation.gov.uk/eut/teec/contents>
- EC (European Commission). (2016a). *Background note on open access to scientific publications and open research data*. https://research-and-innovation.ec.europa.eu/document/download/4bd9ef8e-0101-457d-9fc5-1096c4e8f6f0_en?filename=ec_rtd_background-note-open-access.pdf
- EC (European Commission). (2016b). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the processing of personal data and on the free movement of such data (general data protection regulation)*. OJ L 119/1. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679>

- EC (European Commission). (2017). *Mapping of media literacy practices and actions in EU-28*. European Commission Audio-Visual Observatory. <https://digital-strategy.ec.europa.eu/en/library/reporting-media-literacy-europe>
- EC (European Commission). (2018). *Communication – tackling online disinformation: A European approach*. COM(2018) 236 final. <https://digital-strategy.ec.europa.eu/en/library/communication-tackling-online-disinformation-european-approach>
- EC (European Commission). (2020a). *Assessment of the code of practice on disinformation – achievements and areas for further improvement*. <https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement>
- EC (European Commission). (2020b). *Communication on the European democracy action plan*. COM(2020) 790 final. <https://www.europeansources.info/record/communication-on-the-european-democracy-action-plan>
- EC (European Commission). (2022a). *Contestable and fair markets in the digital sector and amending directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act)*, PE/17/2022/REV/1. Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022. <https://eur-lex.europa.eu/eli/reg/2022/1925/oj>
- EC (European Commission). (2022b). *European Commission digital strategy*. European Commission C(2022) 4388 final. https://commission.europa.eu/publications/european-commission-digital-strategy_en
- EC (European Commission). (2022c). *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending directive 2000/31/EC (Digital Services Act)*. <https://eur-lex.europa.eu/eli/reg/2022/2065/oj>
- EC (European Commission). (2022d). *Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending regulation (EU) 2018/1724 (Data Governance Act)*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022R0868>
- EC (European Commission). (2022e). *The strengthened code of practice on disinformation 2022*. <https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation>
- EC (European Commission). (2023). *Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending regulation (EU) 2017/2394 and directive (EU) 2020/1828 (Data Act)*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202302854
- EC (European Commission). (2024a, May 8). *Commission requests information from X on decreasing content moderation resources under the Digital Services Act*. European Commission News. <https://digital-strategy.ec.europa.eu/en/news/commission-requests-information-x-decreasing-content-moderation-resources-under-digital-services>
- EC (European Commission). (2024b). *Regulation (EU) 2024/1083 establishing a common framework for media service (European Media Freedom Act)*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401083
- EC (European Commission). (2024c). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689
- EC (European Commission). (2024d). *Funded projects in the fight against disinformation*. https://commission.europa.eu/strategy-and-policy/coronavirus-response/fighting-disinformation/funded-projects-fight-against-disinformation_en
- EC (European Commission). (2025). *AI continent action plan* COM(2025)165. https://commission.europa.eu/topics/eu-competitiveness/ai-continent_en
- EC HLG (European Commission High Level Group). (2018). *A multi-dimensional approach to disinformation: Report of the independent high level group on fake news and online disinformation*. <https://op.europa.eu/en/publication-detail/-/publication/6ef4df8b-4cea-11e8-be1d-01aa75ed71a1/language-en>
- Ecker, U., Roozenbeek, J., van der Linden, S., Tay, L. Q., Cook, J., Oreskes, N., & Lewandowsky, S. (2024a). Misinformation poses a bigger threat to democracy than you might think. *Nature*, 630(8015), 29–32. <https://doi.org/10.1038/d41586-024-01587-3>

- Ecker, U., Tay, L. Q., Roozenbeek, J., van der Linden, S., Cook, J., Oreskes, N., & Lewandowsky, S. (2024b). Why misinformation must not be ignored. *PsyArXiv*.
<https://doi.org/10.31234/osf.io/8a6cj>
- Economides, N., & Hermalin, B. E. (2012). The economics of network neutrality. *The RAND Journal of Economics*, 43(4), 602–629. <https://www.jstor.org/stable/41723347>
- Eddy, K. (2024). *Republicans, young adults now nearly as likely to trust info from social media as from national news outlets*. Pew Research Center.
<https://www.pewresearch.org/short-reads/2024/10/16/republicans-young-adults-now-nearly-as-likely-to-trust-info-from-social-media-as-from-national-news-outlets/>
- Edgerly, S. (2022). The head and heart of news avoidance: How attitudes about the news media relate to levels of news consumption. *Journalism*, 23(9), 1828–1845.
<https://doi.org/10.1177/14648849211012922>
- Effoduh, J. O. (2024, April 30). A *Global South perspective on explainable AI*. Carnegie Endowment for International Peace. <https://carnegieendowment.org/2024/04/30/global-south-perspective-on-explainable-ai-pub-92333>
- Egbejule, E. (2025, March 9). Internet shutdowns at record high in Africa as access ‘weaponised’. *The Guardian*. <https://www.theguardian.com/technology/2025/mar/09/internet-shutdowns-record-high-africa-2024-access-weaponised>
- Ehrenfeld, D. (2020). “Sharing a world with others”: Rhetoric’s ecological turn and the transformation of the networked public sphere. *Rhetoric Society Quarterly*, 50(5), 305–320.
<https://doi.org/10.1080/02773945.2020.1813321>
- Eichler, H. (2023). Cross-media and platformised journalism: ARD’s innovation attempts at becoming a multi-platform contributor to society. In M. Puppis, & C. Ali (Eds.), *Public service media’s contribution to society: RIPE@2021* (pp. 265–288). Nordicom, University of Gothenburg. <https://doi.org/10.48335/9789188855756-13>
- Eisenach, J. A. (2015). *The economics of zero rating*. NERA Economic Consulting, US.
<https://tinyurl.com/2f9bxh45>
- Elishar-Malka, V., Ariel, Y., & Weimann, G. (2020). Rethinking political communication in the digital sphere. *The Journal of International Communication*, 26(2), 190–210.
<https://doi.org/10.1080/13216597.2020.1771397>
- Elkin-Koren, N. (2020). Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence. *Big Data & Society*, 7(2), 1–13.
<https://doi.org/10.1177/2053951720932296>
- Elkin-Koren, N., Perel (Filmar), M., & Somech, O. (2024). Unlocking platform data for research. *Proceedings of the TPRC2024, the Research Conference on Communications, Information and Internet Policy*. <https://doi.org/10.2139/ssrn.4733788>
- Elliott, V. (2022, June 23). Meta made millions in ads from networks of fake accounts. *Wired*.
<https://www.wired.com/story/meta-is-making-millions-from-fake-accounts>
- Elliott, V. (2024, September 2). AI-fakes detection is failing voters in the Global South. *Wired*.
<https://www.wired.com/story/generative-ai-detection-gap>
- Elmer, G., & Ward-Kimola, S. (2023). Crowdfunding (as) disinformation: ‘Pitching’ 5G and election fraud campaigns on GoFundMe. *Media, Culture & Society*, 45(3), 578–594.
<https://doi.org/10.1177/01634437221136009>
- Elsom, C. (2020). *Safety without censorship: A better way to tackle online harms*. Centre for Policy Studies, London.
- Engel, F., & Noske-Turner, J. (Eds.). (2018). *Communication in international development: Doing good or looking good?* Routledge. <https://doi.org/10.4324/9780203703977>
- EP (European Parliament). (2024). *EP academic freedom monitor 2023*. EPRS (European Parliamentary Research Service), STOA (Scientific Foresight Unit), PE 757.798.
[https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU\(2024\)757798](https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2024)757798)
- Epstein, B. (2020). Why it is so difficult to regulate disinformation online. In W. L. Bennett, & S. Livingston (Eds.), *The disinformation age* (pp. 190–210). Cambridge University Press.
- Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences of the United States of America*, 112(33), E4512–E4521.
<https://www.pnas.org/doi/full/10.1073/pnas.1419828112>

- ERGA (European Regulators Group for Audiovisual Media Services). (2020). *ERGA report on disinformation: Assessment of the implementation of the code of practice*. <https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf>
- Erickson, J. (2024). Rethinking the filter bubble? Developing a research agenda for the protective filter bubble. *Big Data & Society*, 11(1), 1–4. <https://doi.org/10.1177/20539517241231276>
- Erie, M. S., & Streinz, T. (2021). The Beijing effect: China's 'digital silk road' as transnational data governance. *New York University Journal of International Law and Politics*, 54(1), 1–25. <https://ssrn.com/abstract=3810256>
- Erkmen, Ö. (2023). Data journalism: A systematic literature review. *Journalism Studies*, 25(1), 58–79. <https://doi.org/10.1080/1461670X.2023.2289885>
- Erman, E., & Furendal, M. (2022). The global governance of artificial intelligence: Some normative concerns. *Moral Philosophy and Politics*, 9(2), 267–291. <https://doi.org/10.1515/mopp-2020-0046>
- eSafety Commissioner. (2024a, January 11). *Report reveals the extent of deep cuts to safety staff and gaps in Twitter/X's measures to tackle online hate*. Media Release. <https://www.esafety.gov.au/newsroom/media-releases/report-reveals-the-extent-of-deep-cuts-to-safety-staff-and-gaps-in-twitter/xs-measures-to-tackle-online-hate>
- eSafety Commissioner. (2024b). *Basic online safety expectations: Summary of response from X Corp. (Twitter) to eSafety's transparency notice on online hate*. eSafety Commissioner, Australia. <https://www.esafety.gov.au/newsroom/media-releases/report-reveals-the-extent-of-deep-cuts-to-safety-staff-and-gaps-in-twitter/xs-measures-to-tackle-online-hate>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St Martin's Publishing Group.
- Evanega, S., Lynas, M., Adams, J., & Smolenyak, K. (2022). *Coronavirus misinformation: Quantifying sources and themes in the COVID-19 'infodemic'*. CALS (Cornell Alliance for Science), US. <https://allianceforscience.org/wp-content/uploads/2020/09/Evanega-et-al-Coronavirus-misinformationFINAL.pdf>
- Evans, R., Jackson, D., & Murphy, J. (2023). Google News and machine gatekeepers: Algorithmic personalisation and news diversity in online news search. *Digital Journalism*, 11(9), 1682–1700. <https://doi.org/10.1080/21670811.2022.2055596>
- Evans, S. K., Pearce, K. E., Vitak, J., & Treem, J. W. (2017). Explicating affordances: A conceptual framework for understanding affordances in communication research. *Journal of Computer-Mediated Communication*, 22(1), 35–52. <https://doi.org/10.1111/jcc4.12180>
- Faculty. (2021). *Automated Approaches to Measuring Online Experiences: Executive Summary*. <https://www.bleadon.org.uk/media/other/24400/OFCOM-automated-tooling-report-facultyreport.pdf>
- Faife, C. (2022, February 11). Feds are still using ID.me to scan your face — and human reviewers can't keep up. *The Verge*. <https://www.theverge.com/2022/2/11/22928082/id-me-irs-facial-recognition-overworked-employees>
- Fallis, D. (2015). What is disinformation? *Library Trends*, 63(3), 401–426. <https://muse.jhu.edu/pub/1/article/579342>
- Farahat, M. (2021). *Egypt digital rights landscape report*. Institute for Development Studies, Sussex. https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/15964/Egypt_Report.pdf
- Farkas, J., & Schou, J. (2018). Fake news as a floating signifier: Hegemony, antagonism and the politics of falsehood. *Javnost – The Public*, 25(3), 298–314. <https://doi.org/10.1080/13183222.2018.1463047>
- Farooq, G. (2018). Politics of fake news: How WhatsApp became a potent propaganda tool in India. *Media Watch*, 9(1), 106–117. <https://doi.org/10.15655/mw/2018/v9i1/49279>
- Farrell, H., & Newman, A. (2023). *Underground empire: How america weaponized the world economy*. Allen Lane.
- Fatafta, M. (2024, February 19). *How Meta censors Palestinian voices*. Access Now. <https://www.accessnow.org/publication/how-meta-censors-palestinian-voices>
- Fawzi, N., Steindl, N., Obermaier, M., Prochazka, F., Arlt, D., Blöbaum, B., Dohle, M., Engelke, K. M., Hanitzsch, T., Jakob, N., Jakobs, I., Klawier, T., Post, S., Reinemann, C., Schweiger, W., & Ziegele, M. (2021). Concepts, causes and consequences of trust in news media – A literature review and framework. *Annals of the International Communication Association*, 45(2), 154–174. <https://doi.org/10.1080/23808985.2021.1960181>

- Feathers, T., & Ng, A. (2022, May 26). *Tech industry groups are watering down attempts at privacy regulation, one state at a time*. The Markup. <https://themarkup.org/privacy/2022/05/26/tech-industry-groups-are-watering-down-attempts-at-privacy-regulation-one-state-at-a-time>
- Federal Register (The Daily Journal of the United States Government). (2023). *Executive order on the safe, secure, and trustworthy development and use of artificial intelligence*. <https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence>
- Fendji, J. L. K. E. (2024). From left behind to left out: Generative AI or the next pain of the unconnected. *Harvard Data Science Review* [Special issue 5]. <https://doi.org/10.1162/99608f92.e360e42d>
- Fernandez-Monge, F., Barns, S., Bria, F., & Kattel, R. (2024). Reclaiming data for improved city governance: Barcelona's New Data Deal. *Urban Studies*, 61(7), 1291–1307. <https://doi.org/10.1177/00420980231204835>
- Ferrara, E. (2024a). Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci*, 6(1), 1–15. <https://doi.org/10.3390/sci6010003>
- Ferrara, E. (2024b). GenAI against humanity: Nefarious applications of generative artificial intelligence and large language models. *Journal of Computational Social Science*, 7, 549–569. <https://doi.org/10.1007/s42001-024-00250-1>
- Ferrás-Hernández, X., Nylund, P. A., & Brem, A. (2023). The emergence of dominant designs in artificial intelligence. *California Management Review*, 65(3), 73–91. <https://doi.org/10.1177/00081256231164362>
- Fertmann, M., & Kettemann, M. C. (Eds.). (2021). *Viral information: How states and platforms deal with Covid-19-related disinformation, An exploratory study of 20 countries*. Verlag Hans-Bredow-Institut. <https://doi.org/10.21241/ssoar.75124>
- Feuerriegel, S., DiResta, R., Goldstein, J. A., Kumar, S., Lorenz-Spreen, P., Tomz, M., & Pröllochs, N. (2023). Research can help to tackle AI-generated disinformation. *Nature Human Behaviour*, 7(11), 1818–1821. <https://doi.org/10.1038/s41562-023-01726-2>
- Findlay, M., Ong, L. M., & Zhang, W. (Eds.). (2023). *Elgar companion to regulating AI and big data in emerging economies*. Edward Elgar Publishing.
- Finkel, S. E. (2014). The impact of adult civic education programmes in developing democracies. *Public Administration and Development*, 34(3), 169–181. <https://doi.org/10.1002/pad.1678>
- Finkel, S. E., Neundorff, A., & Rascón Ramírez, E. (2024). Can online civic education induce democratic citizenship? Experimental evidence from a new democracy. *American Journal of Political Science*, 68, 613–630. <https://doi.org/10.1111/ajps.12765>
- Fischer-Lescano, A. (2016). Struggles for a global Internet constitution: Protecting global communication structures against surveillance measures. *Global Constitutionalism*, 5(2), 145–172. <https://doi.org/10.1017/S204538171600006X>
- Fitzgerald, C., Williams, K., & Cross, R. J. (2024). *The state of privacy: How state “privacy” laws fail to protect privacy and how they can do better*. Electronic Privacy Information Center & Public Interest Research Group (PIRG), Education Fund, US. <https://www.documentcloud.org/documents/24400016-state-of-privacy-feb-2024>
- Fletcher, R., & Nielsen, R. K. (2024). *What does the public in six countries think of generative AI in news?* Reuters Institute for the Study of Journalism, University of Oxford. <http://reutersinstitute.politics.ox.ac.uk/what-does-public-six-countries-think-generative-ai-news>
- Flew, T. (2021). *Regulating platforms*. Polity Press.
- Flew, T. (2023). Mediated trust and artificial intelligence. *Emerging Media*, 1(1), 22–29. <https://doi.org/10.1177/27523543231188793>
- Flew, T., & Martin, F. R. (Eds.). (2022). *Digital platform regulation: Global perspectives on internet governance*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-95220-4>
- Flew, T., & Stepnik, A. (2024). The value of news: Aligning economic and social value from an institutional perspective. *Media and Communication*, 12, 1–16. <https://doi.org/10.17645/mac.7462>
- Flew, T., Iosifidis, P., Meese, J., & Stepnik, A. (2023). Digital platforms and the future of news: Regulating publisher–platform relations in Australia and Canada. *Information, Communication & Society*, 27(16), 2727–2743. <https://doi.org/10.1080/1369118X.2023.2291462>

- Floridi, L. (2013). *The ethics of information*. Oxford Academic.
<https://doi.org/10.1093/acprof:oso/9780199641321.001.0001>
- Floridi, L., & Nobre, A. C. (2024). Anthropomorphising machines and computerising minds: The crosswiring of languages between artificial intelligence and brain & cognitive sciences. *Minds and Machines*, 34(5), 1–9. <https://doi.org/10.2139/ssrn.4738331>
- Fonseka, P. (2024, January 5). *Catalysing change: Sri Lanka refines digital literacy definition*. LIRNEasia.
<https://lirneasia.net/2024/01/catalysing-change-sri-lanka-refines-digital-literacy-definition>
- Fontes, C., Hohma, E., Corrigan, C. C., & Lürge, C. (2022). AI-powered public surveillance systems: Why we (might) need them and how we want them. *Technology in Society*, 71, 1–12. <https://doi.org/10.1016/j.techsoc.2022.102137>
- Ford, E. (2019). *El Reto de la Democracia Digital: Hacia una Ciudadanía Interconectada [The challenge of digital democracy: Towards an interconnected citizenship]*. KAS, JNE, ONPE, D&D Internacional. <https://democraciadigital.pe/publicaciones>
- Forum on Information and Democracy. (2023, February). *Pluralism of news and information in curation and indexing algorithms: Policy framework*.
https://informationdemocracy.org/wp-content/uploads/2023/02/Report_Pluralism-in-algorithms.pdf
- Forum on Information and Democracy. (2024a, February). *AI as a public good: Ensuring democratic control of AI in the information space*. <https://informationdemocracy.org/wp-content/uploads/2024/03/ID-AI-as-a-Public-Good-Feb-2024.pdf>
- Forum on Information and Democracy. (2024b). *Data protection authorities stand up to big tech: The importance of democratic institutions to safeguard citizens' privacy rights*. Unpacking Current Developments in the Information Space, Insight Nr. 1.
<https://informationdemocracy.org/wp-content/uploads/2024/09/FID-Insight-Nr-1-Data-protection-authorities-and-AI.pdf>
- Forum on Information and Democracy. (2024c, June 14). *The EU's Digital Services Act must provide researchers access to VLOPs' experimental protocols*. Press Release on the paper of Alistair Knott, Dino Predreschi, Jonathan Stray & Stuart Russell.
<https://informationdemocracy.org/2024/06/14/the-forum-members-of-its-working-group-and-researchers-call-for-the-dsa-to-provide-researchers-access-to-conduct-experimental-evaluations-of-vlops>
- Forum on Information and Democracy. (2024d). *US Supreme Court decision on social media regulation, uncertainty remains!* Unpacking Current Developments in the Information Space, Insight Nr. 2. <https://informationdemocracy.org/wp-content/uploads/2024/09/FID-Insight-Nr-2-US-Supreme-Court-Decision-on-social-media-regulation.pdf>
- Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings, 1972–1977*. Knopf Doubleday.
- Franco, S. F., Graña, J. M., & Rikap, C. (2024). Dependency in the digital age? The experience of Mercado Libre in Latin America. *Development and Change*, 55(3), 429–464.
<https://ideas.repec.org/a/bla/devchg/v55y2024i3p429-464.html>
- François, C. (2019). *Actors, behaviors, content: A disinformation ABC: Highlighting three vectors of viral deception to guide industry & regulatory responses*. TransAtlantic Working Group, Graphika and Berkman Klein Center for Internet & Society, Harvard University.
<https://docs.house.gov/meetings/SY/SY21/20190926/109980/HHRG-116-SY21-Wstate-FrancoisC-20190926-SD001.pdf>
- Franks, M. A. (2021). Beyond the public square: Imagining digital democracy. *Yale Law Journal*, 131(2021-22), 427–453.
<https://www.yalelawjournal.org/forum/beyond-the-public-square-imagining-digital-democracy>
- Fraser, N. (1992). Rethinking the public sphere: A contribution to the critique of actually existing democracy. In C. Calhoun (Ed.), *Habermas and the public sphere* (pp. 109–142). MIT Press.
- Fraser, N. (2022). *Cannibal capitalism: How our system is devouring democracy, care, and the planet – and what we can do about it*. Verso.
- Fraser, N., & Nash, K. (Eds.). (2014). *Transnationalizing the public sphere*. Polity Press.
- Frau-Meigs, D. (2022). How disinformation reshaped the relationship between journalism and media and information literacy (MIL): Old and new perspectives revisited. *Digital Journalism*, 10(5), 912–922. <https://doi.org/10.1080/21670811.2022.2081863>

- Frau-Meigs, D. (2024a). Algorithm literacy as a subset of media and information literacy: Competences and design considerations. *Digital*, 4(2), 512–528. <https://doi.org/10.3390/digital4020026>
- Frau-Meigs, D. (2024b, March). *On AI pedagogy*. Presentation to VOICES – European Festival of Journalism and Media Literacy. https://www.pedagogie.ac-aix-marseille.fr/upload/docs/application/pdf/2024-05/lbn_25_2024-05-12_20-03-43_354.pdf
- Frau-Meigs, D. (2024c). *User empowerment through media and information literacy responses to the evolution of generative artificial intelligence (GAI)*. Report prepared for UNESCO, Sorbonne Nouvelle University. <https://unesdoc.unesco.org/ark:/48223/pf0000388547>
- Frau-Meigs, D., & Corbu, N. (Eds.). (2024). *Disinformation debunked: Building resilience through media and information literacy*. Routledge. <https://doi.org/10.4324/9781003387404>
- Frau-Meigs, D., Velez, I., & Michel, J. F. (Eds.). (2017). *Public policies in media and information literacy in Europe: Cross-country comparisons*. Routledge. <https://doi.org/10.4324/9781315628851>
- Fredheim, R., Bay, S., Dek, A., Stolze, M., & Haiduchyk, T. (2023). *Social media manipulation 2022/2023: Assessing the ability of social media companies to combat platform manipulation*. NATO Strategic Communications Centre of Excellence. <https://stratcomcoe.org/publications/social-media-manipulation-20222023-assessing-the-ability-of-social-media-companies-to-combat-platform-manipulation/272>
- Freedman, D. (2014). The puzzle of media power: Notes toward a materialist approach. *International Journal of Communication*, 8(2014), 319–334. <https://ijoc.org/index.php/ijoc/article/view/2081>
- Freedom House. (2024). *The struggle for trust online*. Freedom on the Net 2024. <https://freedomhouse.org/report/freedom-net/2024/struggle-trust-online>
- Freedom Online Coalition. (n.d.). *Aims and priorities*. Retrieved September 2, 2025 from <https://freedomonlinecoalition.com/aims-and-priorities>
- Freelon, D., & Wells, C. (2020). Disinformation as political communication. *Political Communication*, 37(2), 145–156. <https://doi.org/10.1080/10584609.2020.1723755>
- Freelon, D., Bossetta, M., Wells, C., Lukito, J., Xia, Y., & Adams, K. (2022). Black trolls matter: Racial and ideological asymmetries in social media disinformation. *Social Science Computer Review*, 40(3), 560–578. <https://doi.org/10.1177/0894439320914853>
- Freire, P. (1974). *Education for critical consciousness*. Continuum.
- Frieden, R. (2014). *Ex ante versus ex post approaches to network neutrality: A cost benefit analysis*. <http://dx.doi.org/10.2139/ssrn.2493945>
- Friesem, Y., Raman, U., Kanižaj, I., & Choi, G. Y. (Eds.). (2022). *The Routledge handbook of media education futures post-pandemic*. Routledge. <https://doi.org/10.4324/9781003283737>
- Frislich, L., & Humprecht, E. (2021). *Trust, democratic resilience, and the infodemic* [Policy paper]. Israel Public Policy Institute, Heinrich Böll Stiftung; Embassy of the Federal Republic of Germany. <https://tinyurl.com/4wphnjp2>
- FTC (Federal Trade Commission). (2024, January 25). FTC launches inquiry into generative AI investments and partnerships [Press release]. <https://www.ftc.gov/news-events/news/press-releases/2024/01/ftc-launches-inquiry-generative-ai-investments-partnerships>
- Fuchs, C. (2023). *Digital democracy and the digital public sphere: Media communication and society volume six*. Routledge. <https://doi.org/10.4324/9781003331087>
- Full Fact. (2025, May 11). *In a 'misinformation' age, we need to know what the fact is going on*. Full Fact. <https://fullfact.org/politics/in-a-misinformation-age-we-need-to-know-what-the-fact-is-going-on/>
- Fung, A., & Wright, E. O. (Eds.). (2003). *Deepening democracy: Institutional innovations in empowered participatory governance*. Verso.
- Funk, A., Shahbaz, A., & Vesteinsson, K. (2023). *Freedom on the net 2023: The repressive power of artificial intelligence*. Freedom House. <https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence>
- Furstenberg, S., Prelic, T., & Heathershaw, J. (2020). *The internationalization of universities and the repression of academic freedom*. Freedom House. <https://freedomhouse.org/report/special-report/2020/internationalization-universities-and-repression-academic-freedom>
- G20. (2019). *G20 ministerial statement on trade and digital economy, annex G20 AI principles*. <https://wp.oecd.ai/app/uploads/2021/06/G20-AI-Principles.pdf>

- G7. (2023). *Hiroshima process international guiding principles for organizations developing advanced AI systems*. G7 2023 Hiroshima Summit. <https://www.mofa.go.jp/files/100573471.pdf>
- Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000233231>
- Galantino, S. (2023). How will the EU Digital Services Act affect the regulation of disinformation? *SCRIPTed: A Journal of Law, Technology and Society*, 20(1), 89–129. https://script-ed.org/wp-content/uploads/2023/02/Sharon-Galantino_February-2023.pdf?d=05112024
- Galič, M. (2022). Smart cities as ‘big brother only to the masses’: The limits of personal privacy and personal surveillance. *Surveillance & Society*, 20(3), 306–311. <https://doi.org/10.2139/ssrn.4228444>
- Gallagher, A., O’Connor, C., & Visser, F. (2023). *Uisce Faoi Thalamh: Summary report – an investigation into the online mis- and disinformation ecosystem in Ireland*. ISD (Institute for Strategic Dialogue). <https://www.isdglobal.org/isd-publications/uisce-faoi-thalamh-summary-report/>
- Gallagher, M. (2023). Gendered disinformation and platform accountability. In M. Gallagher, & A. Vega Montiel (Eds.), *The handbook of gender, communication, and women’s human rights* (pp. 53–69). John Wiley & Sons.
- Galli, F., Loreggia, A., & Sartor, G. (2023). The regulation of content moderation. In D. Moura Vicente, S. de Vasconcelos Casimiro, & C. Chen (Eds.), *The legal challenges of the fourth industrial revolution: The European Union’s digital strategy* (pp. 63–87). Springer.
- Garcia Ramirez, D. (2021). Journalism in the attention economy – The relation between digital platforms and news organisations. *Brazilian Journalism Research*, 17(1), 1–24. <https://doi.org/10.25200/BJR.v17n1.2021.1332>
- Garcia Santamaria, S., & Salojärvi, V. (2020). Media in authoritarian contexts: A logics approach to journalistic professional resistance in Cuba and Venezuela. In X. Orchard, S. Garcia Santamaria, J. Brambila, & J. Lugo-Ocando (Eds.), *Media & governance in Latin America: Towards plurality of voices* (pp. 97–116). Peter Lang.
- Gates, B. (2023, March 21). *The age of AI has begun* [Blog post]. GatesNotes. <https://www.gatesnotes.com/The-Age-of-AI-Has-Begun>
- Gautier, A., & Lamesch, J. (2021). Mergers in the digital economy. *Information Economics and Policy*, 54, 1–15. <https://doi.org/10.1016/j.infoecopol.2020.100890>
- Geiger, C. (2024). Elaborating a human rights-friendly copyright framework for generative AI. *IIC – International Review of Intellectual Property and Competition Law*, 55, 1129–1165. <https://doi.org/10.1007/s40319-024-01481-5>
- Geiß, S., Magin, M., Jürgens, P., & Stark, B. (2021). Loopholes in the echo chambers: How the echo chamber metaphor oversimplifies the effects of information gateways on opinion expression. *Digital Journalism*, 9(5), 660–686. <https://doi.org/10.1080/21670811.2021.1873811>
- Georgiou, M. (2023). *Being human in digital cities*. Polity Press.
- Gerbaudo, P. (2022). Theorizing reactive democracy: The social media public sphere, online crowds, and the plebiscitary logic of online reactions. *Democratic Theory*, 9(2), 120–138. <https://doi.org/10.3167/dt.2022.090207>
- Gerbner, G., Gross, L., Signorielli, N., & Morgan, M. (1980). Aging with television: Images on television drama and conceptions of social reality. *Journal of Communication*, 30(1), 37–47. <https://eric.ed.gov/?id=EJ221287>
- Germain, T. (2024, January 9). *X bans and then unbans journalists and podcasters in Twitter’s latest free speech massacre*. Gizmodo. <https://gizmodo.com/elon-musk-x-twitter-journalists-banning-spree-1851151593>
- German Federal Ministry of Education and Research. (2021, February 24). Guideline for the federal-state initiative to promote artificial intelligence in higher education. *Federal Gazette*. https://www.bmfr.bund.de/SharedDocs/Bekanntmachungen/DE/2021/02/3409_bekanntmachung.html
- Gerpott, T. J. (2018). Zero-rating arrangements of mobile internet access service providers – An analysis of main factors shaping the need for regulatory interventions. *Telecommunications Policy*, 42(6), 489–500. <https://doi.org/10.1016/j.telpol.2018.03.003>

- Ghosal, S. S., Chakraborty, S., Geiping, J., Huang, F., Manocha, D., & Bedi, A. S. (2023). Towards possibilities & impossibilities of AI-generated text detection: A survey. *ArXiv*. <https://doi.org/10.48550/arXiv.2310.15264>
- Ghose, R. (2001). Use of information technology for community empowerment: Transforming geographic information systems into community information systems. *Transactions in GIS*, 5(2), 141–163. <https://doi.org/10.1111/1467-9671.00073>
- Ghosh, D., & Scott, B. (2018). *Digital deceit: The technologies behind precision propaganda on the internet*. New America, Public Interest Technology, & Harvard Kennedy School Shorenstein Center on Media, Politics and Public Policy. <http://newamerica.org/pit/policy-papers/digitaldeceit/>
- Giblin, R., & Doctorow, C. (2022). *Chokepoint capitalism: How big tech and big content captured creative labor markets and how we'll win them back*. Beacon Press.
- Gibson, A. D., Docherty, N., & Gillespie, T. (2023). Health and toxicity in content moderation: The discursive work of justification. *Information, Communication & Society*, 27(7), 1441–1457. <https://doi.org/10.1080/1369118X.2023.2291456>
- Gil de Zúñiga, H., & Goyanes, M. (2023). Fueling civil disobedience in democracy: WhatsApp news use, political knowledge, and illegal political protest. *New Media & Society*, 25(12), 3500–3520. <https://doi.org/10.1177/14614448211047850>
- Gil de Zúñiga, H., Borah, P., & Goyanes, M. (2021). How do people learn about politics when inadvertently exposed to news? Incidental news paradoxical direct and indirect effects on political knowledge. *Computers in Human Behavior*, 121, 1–9. <https://doi.org/10.1016/j.chb.2021.106803>
- Gil de Zúñiga, H., Goyanes, M., & Durotoye, T. (2023). A scholarly definition of artificial intelligence (AI): Advancing AI as a conceptual framework in communication research. *Political Communication*, 41(2), 317–334. <https://doi.org/10.1080/10584609.2023.2290497>
- Gillespie, T. (2010). The politics of ‘platforms’. *New Media & Society*, 12(3), 347–364. <https://doi.org/10.1177/1461444809342738>
- Gillespie, T. (2020). Content moderation, AI, and the question of scale. *Big Data & Society*, 7(2), 1–5. <https://doi.org/10.1177/2053951720943234>
- Gillespie, T., Aufderheide, P., Carmi, E., Gerrard, Y., Gorwa, R., Matamoros-Fernández, A., Roberts, S. T., Sinnreich, A., & Myers West, S. (2023). Expanding the debate about content moderation: Scholarly research agendas for the coming policy debates. *Internet Policy Review*, 9(4), 1–19. <https://ssrn.com/abstract=4459448>
- Gillwald, A. (2024). *Harnessing global cooperation: Africa's role in shaping the digital future after the Global Digital Compact*. Research ICT Africa. <https://mailchi.mp/researchictafrica/africas-role-in-shaping-the-digital-future-after-the-global-digital-compact?e=43ad1c2882>
- Gillwald, A., & Wavre, V. (2024). Rerouting geopolitics: Narratives and the political power of communication. In C. Padovani, V. Wavre, A. Hintz, G. Goggin, & P. Iosifidis (Eds.), *Global communication governance at the crossroads* (pp. 21–35). Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-29616-1>
- Gillwald, A., & Weleilakeba, S. (2024). *From the digital inequality paradox to marginalisation through digitalisation: GISWatch 2024 special edition reports*. Association for Progressive Communications. <https://giswatch.org/sites/default/files/GW2024-Gillwald.pdf>
- Gitelman, L. (2006). *Always already new: Media, history, and the data of culture*. MIT Press. <https://doi.org/10.7551/mitpress/1208.001.0001>
- Gitelman, L. (Ed.) (2013). ‘Raw data’ is an oxymoron. MIT Press.
- Giusti, S., & Piras, E. (Eds.). (2021). *Democracy and fake news: Information manipulation and post-truth politics*. Routledge. <https://doi.org/10.4324/9781003037385>
- Glasberg, D. S., & Shannon, D. (2010). *Political sociology: Oppression, resistance, and the state*. Sage.
- Glimmerveen, L., Ybema, S., & Nies, H. (2022). Who participates in public participation? The exclusionary effects of inclusionary efforts. *Administration & Society*, 54(4), 543–574. <https://doi.org/10.1177/00953997211034137>
- Global Kids Online. (2019). *Global kids online: Comparative report*. UNICEF Office of Research – Innocenti. <https://www.unicef.org/innocenti/media/7011/file/GKO-Comparative-Report-2019.pdf>

- Global Witness. (2023, November 30). *How big tech platforms are neglecting their non-English language users*. <https://www.globalwitness.org/en/campaigns/digital-threats/how-big-tech-platforms-are-neglecting-their-non-english-language-users>
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harvard University Press.
- Goldman Sachs. (2024, May 14). *AI is poised to drive 160% increase in data center power demand*. Goldman Sachs. <https://www.goldmansachs.com/insights/articles/AI-poised-to-drive-160-increase-in-power-demand>
- Goldstein, J. A., Chao, J., Grossman, S., Stamos, A., & Tomz, M. (2024). How persuasive is AI-generated propaganda? *Proceedings of the National Academy of Sciences Nexus*, 3(2), 1–7. <https://doi.org/10.1093/pnasnexus/pgae034>
- Gollwitzer, A., Martel, C., Brady, W. J., Pärnamets, P., Freedman, I. G., Knowles, E. D., & Van Bavel, J. J. (2020). Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic. *Nature Human Behaviour*, 4(11), 1186–1197. <https://doi.org/10.1038/s41562-020-00977-7>
- Gonçalves, J., Weber, I., Masullo, G. M., Torres da Silva, M., & Hofhuis, J. (2023). Common sense or censorship: How algorithmic moderators and message type influence perceptions of online content deletion. *New Media & Society*, 25(10), 2595–2617. <https://doi.org/10.1177/14614448211032310>
- González-Bailón, S., Lazer, D., Barberá, P., Zhang, M., et al. (2023). Asymmetric ideological segregation in exposure to political news on Facebook. *Science*, 381(6656), 392–398. <https://doi.org/10.1126/science.ade7138>
- González-Bailón, S., Lazer, D., Barberá, P., Godel, W., Allcott, H., Brown, T., Crespo-Tenorio, A., Freelon, D., Gentzkow, M., Guess, A. M., Iyengar, S., Kim, Y. M., Malhotra, N., Moehler, D., Nyhan, B., Pan, J., Rivera, C. V., Settle, J., Thorson, E., ... Tucker, J. A. (2024). The diffusion and reach of (mis)information on Facebook during the U.S. 2020 election. *Sociological Science*, 11, 1124–1146. <https://doi.org/10.15195/v11.a41>
- Google. (2022, November 21). *Our commitment to climate-conscious data center cooling*. <https://blog.google/outreach-initiatives/sustainability/our-commitment-to-climate-conscious-data-center-cooling>
- Gordon-Tapiero, A., Ohm, P., & Ramaswami, A. (2023, November). Fact and friction: A case study in the fight against false news. *UC Davis Law Review*, 57, 171–251. <https://lawreview.law.ucdavis.edu/archives/57/1/fact-and-friction-case-study-fight-against-false-news>
- Gorwa, R., & Guilbeault, D. (2020). Unpacking the social media bot: A typology to guide research and policy. *Policy & Internet*, 12(2), 225–248. <https://doi.org/10.1002/poi3.184>
- Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1), 1–15. <https://doi.org/10.1177/2053951719897945>
- Gottfredsen, S. G., & Dowling, K. (2024, July 9). Meta is getting rid of CrowdTangle – And its replacement isn't as transparent or accessible. *Columbia Journalism Review*. https://www.cjr.org/tow_center/meta-is-getting-rid-of-crowdtangle.php
- Government of Bangladesh. (2023). *Cyber Security Act, 2023*. <https://www.refworld.org/legal/legislation/natlegbod/2023/en/148277>
- Government of Brazil. (2018). *Lei Geral de Proteção de Dados Pessoais (LGPD) [General law on the protection of personal data]*. <https://www.mpf.mp.br/servicos/lgpd/o-que-e-a-lgpd>
- Government of Brazil. (2024). *Projeto de Lei 2338/2023, Dispõe sobre o uso da Inteligência Artificial [Bill 2338/2023, provides for the use of artificial intelligence]*. <https://artificialintelligenceact.com/brazil-ai-act/>
- Government of France. (2018). *LOI organique n° 2018–1201 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information (1) [Organic Law No. 2018–1201 of 22 December 2018 regarding the fight against information manipulation]*. <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000037847556>
- Government of Germany. (2017). *NetzDG – Netzdurchsetzungsgesetz [Network enforcement act]*. <https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html>
- Government of India. (2023a). *The Digital Personal Data Protection Act, 2023*. Ministry of Electronics & Information Technology. <https://tinyurl.com/3n57zbjb>

- Government of India. (2023b). *The information technology (intermediary guidelines and digital media ethics code) rules, 2021* [updated as on 6.4.2023]. Ministry of Electronics and Information Technology. <https://tinyurl.com/mrymsnzs>
- Graham, M., & Dittus, M. (2022). *Geographies of digital exclusion: Data and inequality*. Pluto Press.
- Graham, M., & Ferrari, F. (Eds.). (2022). *Digital work in the planetary market*. MIT Press.
- Graham, T. (2024, October 3). Is big tech harming society? To find out, we need research – but it's being manipulated by big tech itself. *The Conversation*. <https://theconversation.com/is-big-tech-harming-society-to-find-out-we-need-research-but-its-being-manipulated-by-big-tech-itself-240110>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Graves, L., & Cherubini, F. (2016). *The rise of fact-checking sites in Europe*. Reuters Institute for the Study of Journalism, University of Oxford. <https://doi.org/10.60625/risj-tdn4-p140>
- Green, B. (2022). Escaping the impossibility of fairness: From formal to substantive algorithmic fairness. *Philosophy & Technology*, 35(4), 1–32. <https://doi.org/10.1007/s13347-022-00584-6>
- Green, B., & Kak, A. (2021, June 15). The false comfort of human oversight as an antidote to AI harm. *Slate*. <https://ssrn.com/abstract=4046163>
- Greenberg, W. (2022, August 31). *How ad tech became cop spy tech*. Electronic Frontier Foundation. <https://www.eff.org/deeplinks/2022/08/how-ad-tech-became-cop-spy-tech>
- Greene, L. (2018). *Silicon states: The Power and politics of big tech and what it means for our future*. Counterpoint.
- Greenstein, S., Peitz, M., & Valletti, T. (2016). Net neutrality: A fast lane to understanding the trade-offs. *Journal of Economic Perspectives*, 30(2), 127–150. <https://doi.org/10.1257/jep.30.2.127>
- Gritsenko, D., Markham, A., Pötzsch, H., & Wijermars, M. (2022). Algorithms, contexts, governance: An introduction to the special issue. *New Media & Society*, 24(4), 835–844. <https://doi.org/10.1177/14614448221079037>
- Grizzle, A., Wilson, C., Tuazon, R., Cheung, C. K., Lau, J., Fischer, R., Gordon, D., Akyempong, K., Singh, J., Carr, P. R., Stewart, K., Tayie, S., Suraj, O., Jaakkola, M., Thésée, G., Gulston, C., Andzongo Menyeng, B. P., & Zibi Fama, P. A. (2021). *Media and information literate citizens: Think critically, click wisely!* UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000377068>
- Grover, R., & Baik, J. (2024). Platforms as templates: Emerging datafication dynamics in digital news outlets' datawalls. *The Information Society*, 40(4), 260–276. <https://doi.org/10.1080/01972243.2024.2350015>
- Grover, R., Widera, J., & Ananny, M. (2025). Data disaffection: Toward a relational and affective understanding of datafication. *New Media & Society*, 1–16. Online First. <https://doi.org/10.1177/14614448241310901>
- Guallar, J., Codina, L., Freixa, P., & Pérez-Montoro, M. (2022). Disinformation, hoaxes, curation and verification: Review of studies in Ibero-America 2017–2020. *Online Media and Global Communication*, 1(3), 648–668. <https://doi.org/10.1515/omgc-2022-0055>
- Guess, A. M., Malhotra, N., Pan, J., Barbera, P., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Kennedy, E., Kim, Y. M., Lazer, D., Moehler, D., Nyhan, B., Velasco Rivera, C., Settle, J., Thomas, D. R., ... Tucker, J. A. (2023a). How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, 381(6656), 398–404. <https://www.science.org/doi/10.1126/science.abp9364>
- Guess, A. M., Malhotra, N., Pan, J., Barbera, P., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Kennedy, E., Kim, Y. M., Lazer, D., Moehler, D., Nyhan, B., Velasco Rivera, C., Settle, J., Thomas, D. R., ... Tucker, J. A. (2023b). Reshares on social media amplify political news but do not detectably affect beliefs or opinions. *Science*, 381(6656), 404–408. <https://www.science.org/doi/10.1126/science.add8424>

- Guess, A. M., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5(1), 1–8.
<https://doi.org/10.1126/sciadv.aau4586>
- Gulati, R. (2023). Meta's oversight board and transnational hybrid adjudication – What consequences for international law? *German Law Journal*, 24(3), 473–493.
<https://doi.org/10.1017/glj.2023.34>
- Gunkel, D. J. (Ed.). (2024). *Handbook on the ethics of artificial intelligence*. Edward Elgar Publishing. <https://doi.org/10.4337/9781803926728>
- Gunton, K. (2022). The impact of the internet and social media platforms on radicalisation to terrorism and violent extremism. In R. Montasari, F. Carroll, I. Mitchell, S. Hara, & R. Bolton-King (Eds.), *Privacy, security and forensics in the internet of things (IoT)* (pp. 167–177). Springer. https://doi.org/10.1007/978-3-030-91218-5_8
- Guo, Z., Schlichtkrull, M., & Vlachos, A. (2022). A survey on automated fact-checking. *Transactions of the Association for Computational Linguistics*, 10, 178–206.
https://doi.org/10.1162/tacl_a_00454
- Gurie, S., & Treisman, D. (2019). Informational autocrats. *The Journal of Economic Perspectives*, 33(4), 100–127. <https://www.jstor.org/stable/26796838>
- Gurie, S., Henry, E., Marquis, T., & Zhuravskaya, E. (2023). *Curtailling false news, amplifying truth*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4616553
- Gurumurthy, A., & Bharthur, D. (2023). *Reframing AI governance through a political economy lens*. IT for Change.
<https://itforchange.net/index.php/reframing-ai-governance-through-a-political-economy-lens>
- Gurumurthy, A., & Chami, N. (2019). *The wicked problem of AI governance*. IT for Change; Friedrich Ebert Stiftung. <https://tinyurl.com/3fmw8p26>
- Gurumurthy, A., & Chami, N. (2022). *Governing the resource of data: To what end and for whom? Conceptual building blocks of a semi-commons approach*. IDFC Institute Working Paper No. 23. Data Governance Network, International Development Finance Club.
<https://itforchange.net/sites/default/files/1741/WP23-Governing-the-Resource-of-Data-AG-NC.pdf>
- Gurumurthy, A., & Chami, N. (2024). Justice in the post-public sphere: The new challenge for global governance. In C. Padovani, V. Wavre, A. Hintz, G. Goggin, & P. Iosifidis (Eds.), *Global communication governance at the crossroads* (pp. 111–127). Palgrave Macmillan.
https://doi.org/10.1007/978-3-031-29616-1_7
- Gutiérrez-Coba, L. M., Coba-Gutiérrez, P., & Gómez-Díaz, J. A. (2020). Fake news about Covid-19: A comparative analysis of six Iberoamerican countries. *Revista Latina de Comunicación Social*, 78, 237–264. <https://doi.org/10.4185/RLCS-2020-1476>
- Ha, S., Park, Y., Kim, J., & Kim, S. (2023). Research trends of digital platforms: A survey of the literature from 2018 to 2021. *Telecommunications Policy*, 47(8), 1–16.
<https://doi.org/10.1016/j.telpol.2023.102543>
- Haas, J., & Kettemann, M. (2024). *Platform and content governance in times of crisis*. OSCE (Organization for Security and Co-operation in Europe).
<https://www.osce.org/files/f/documents/2/4/564961.pdf>
- Habermas, J. (2015). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. John Wiley & Sons. (Original work published in 1962)
- Habermas, J. (2022). Reflections and hypotheses on a further structural transformation of the political public sphere. *Theory, Culture & Society*, 39(4), 145–171.
<https://doi.org/10.1177/02632764221112341>
- Hadi Mogavi, R., Deng, C., Juho Kim, J., Zhou, P., Kwon, Y. D., Metwally, A. H. S., Tlili, A., Bassanelli, S., Bucchiarone, A., Gujar, S., Nacke, L. E., & Hui, P. (2024). ChatGPT in education: A blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. *Computers in Human Behavior: Artificial Humans*, 2(1), 1–20.
<https://doi.org/10.1016/j.chbah.2023.100027>
- Haenig, M. A., & Ji, X. (2024). A tale of two Southeast Asian states: Media governance and authoritarian regimes in Singapore and Vietnam. *Asian Review of Political Economy*, 3(4), 1–23. <https://doi.org/10.1007/s44216-024-00024-6>
- Haider, J., & Sundin, O. (2022). *Paradoxes of media and information literacy: The crisis of information*. Routledge. <https://doi.org/10.4324/9781003163237>

- Hall, P., & Ellis, D. (2023). A systematic review of socio-technical gender bias in AI algorithms. *Online Information Review*, 47(7), 1264–1279. <https://doi.org/10.1108/OIR-08-2021-0452>
- Hallin, D. C., & Mancini, P. (2004). *Comparing media systems: Three models of media and politics*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511790867>
- Hallin, D. C., & Mancini, P. (Eds.). (2012). *Comparing media systems beyond the western world*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139005098>
- Halpern, H., Brown, G., Kagan, O., Li, L., & Avramenko, A. (2024, June 14). *Google accused of privacy breaches over alleged Chrome tracking*. GRC World Forums. <https://www.grcworldforums.com/risk/google-accused-of-privacy-breaches-over-alleged-chrome-tracking/9675.article>
- Hameleers, M., & Marquart, F. (2023). It's nothing but a deepfake! The effects of misinformation and deepfake labels delegitimizing an authentic political speech. *International Journal of Communication*, 17, 6291–6307. <https://ijoc.org/index.php/ijoc/article/view/20777>
- Hamelink, C. J. (2023). *Communication and human rights: Towards communicative justice*. Polity.
- Han, T. A., Lenaerts, T., Santos, F. C., & Pereira, L. M. (2022). Voluntary safety commitments provide an escape from over-regulation in AI development. *Technology in Society*, 68, 1–14. <https://doi.org/10.1016/j.techsoc.2021.101843>
- Hancock, E. (2024, March 6). *The EU's uphill battle against big tech power*. Politico. <https://www.politico.eu/article/the-eus-uphill-battle-against-big-tech-power>
- Hanitzsch, T., & Vos, T. P. (2018). Journalism beyond democracy: A new look into journalistic roles in political and everyday life. *Journalism*, 19(2), 146–164. <https://doi.org/10.1177/1464884916673386>
- Hanitzsch, T., Van Dalen, A., & Steindl, N. (2018). Caught in the nexus: A comparative and longitudinal analysis of public trust in the press. *The International Journal of Press/Politics*, 23(1), 3–23. <https://doi.org/10.1177/1940161217740695>
- Hao, K. (2020, December 4). We read the paper that forced Timnit Gebru out of Google. Here's what it says. *MIT Technology Review*. <https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru>
- Hao, K. (2022, April 22). A new vision of artificial intelligence for the people. *MIT Technology Review*. <https://www.technologyreview.com/2022/04/22/1050394/artificial-intelligence-for-the-people>
- Hao, X., Demir, E., & Eysers, D. (2024). Exploring collaborative decision-making: A quasi-experimental study of human and generative AI interaction. *Technology in Society*, 78, 1–22. <https://doi.org/10.1016/j.techsoc.2024.102662>
- Harbath, K. (2023, December 19). *Different approaches to counting elections*. Anchor Change. <https://anchorchange.substack.com/p/different-approaches-to-counting>
- Hardy, C., & Williams, S. (2008). E-government policy and practice: A theoretical and empirical exploration of public e-procurement. *Government Information Quarterly*, 25(2), 155–180. <https://doi.org/10.1016/j.giq.2007.02.003>
- Hargittai, E. (Ed.). (2021). *Handbook of digital inequality*. Edward Elgar Publishing. <https://doi.org/10.4337/9781788116572>
- Harlow, S. (2017). *Liberation technology in El Salvador: Re-appropriating social media among alternative media projects*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-48039-8>
- Harlow, S. (2022). A new people's press? Understanding digital-native news sites in Latin America as alternative media. *Digital Journalism*, 10(8), 1322–1341. <https://doi.org/10.1080/21670811.2021.1907204>
- Harrington, E., & Vermeulen, M. (2024). *External researcher access to closed foundation models: State of the field and options for improvement*. AWO supported by the Mozilla Foundation. <https://blog.mozilla.org/wp-content/blogs.dir/278/files/2024/10/External-researcher-access-to-closed-foundation-models.pdf>
- Harris, B. (2023, May 2). Google draws backlash from Brazil with lobbying against 'fake news' bill. *The Financial Times*. <https://www.ft.com/content/8a4bb131-e792-491d-a348-9dee4de41ce2>
- Harris, J. (2023, May 22). 'There was all sorts of toxic behaviour': Timnit Gebru on her sacking by Google, AI's dangers and big tech's biases. *The Guardian*. <https://www.theguardian.com/lifeandstyle/2023/may/22/there-was-all-sorts-of-toxic-behaviour-timnit-gebru-on-her-sacking-by-google-ais-dangers-and-big-techs-biases>

- Harsin, J. (2024). Three critiques of disinformation (for-hire) scholarship: Definitional vortexes, disciplinary unneighborliness, and cryptonormativity. *Social Media + Society*, 10(1), 1–12. <https://doi.org/10.1177/20563051231224732>
- Hartley, J. M., Petre, C., Bengtsson, M., & Kammer, A. (2023). Autonomies and dependencies: Shifting configurations of power in the platformization of news. *Digital Journalism*, 11(8), 1375–1390. <https://doi.org/10.1080/21670811.2023.2257759>
- Harvie, D., Lightfoot, G., Lilley, S., & Weir, K. (2013). Publisher, be damned! From price gouging to the open road. *Prometheus*, 31(3), 229–239. <https://doi.org/10.1080/08109028.2014.891710>
- Harwell, D. (2019, July 7). FBI, ICE find state driver's license photos are a gold mine for facial-recognition searches. *The Washington Post*. <https://www.washingtonpost.com/technology/2019/07/07/fbi-ice-find-state-drivers-license-photos-are-gold-mine-facial-recognition-searches>
- Harwell, D. (2021, February 26). ICE investigators used a private utility database covering millions to pursue immigration violations. *The Washington Post*. <https://www.washingtonpost.com/technology/2021/02/26/ice-private-utility-data>
- Hase, V., Boczek, K., & Scharrow, M. (2023). Adapting to affordances and audiences? A cross-platform, multi-modal analysis of the platformization of news on Facebook, Instagram, TikTok, and Twitter. *Digital Journalism*, 11(8), 1499–1520. <https://doi.org/10.1080/21670811.2022.2128389>
- Hasimi, L., & Poniszewska-Maraña, A. (2024). Detection of disinformation and content filtering using machine learning: Implications to human rights and freedom of speech. *ROMCIR 2024: The 4th Workshop on Reducing Online Misinformation through Credible Information Retrieval (held as part of ECIR 2024, the 46th European Conference on Information Retrieval)*. <https://ceur-ws.org/Vol-3677/paper6.pdf>
- Hatfield, A., Meghiani, K., Williams-Bellamy, J., Saiyid, D., & Zhou, J. (2023). *Mapping and recommendations for UNDP efforts to fight against mis- and disinformation*. Columbia SIPA and UNDP Accelerator Labs. <https://www.undp.org/acceleratorlabs/publications/mapping-and-recommendations-undp-efforts-fight-against-mis-and-disinformation>
- He, A. (2023). *State-centric data governance in China*. CIGI Papers No. 282. CIGI (Centre for International Governance Innovation), Canada. <https://tinyurl.com/4kytsjs7>
- Heeks, R. (2022). Digital inequality beyond the digital divide: Conceptualizing adverse digital incorporation in the global South. *Information Technology for Development*, 28(4), 688–704. <https://doi.org/10.1080/02681102.2022.2068492>
- Heitz, L., Rozgonyi, K., & Kostic, B. (2021). AI in content curation and media pluralism. In D. Wagner, & J. Haas (Eds.), *Spotlight on artificial intelligence and freedom of expression – a policy manual* (pp. 56–70). OSCE (Organization for Security and Co-operation in Europe). <https://www.osce.org/representative-on-freedom-of-media/456319>
- Helberger, N., Karppinen, K., & D'Acunto, L. (2018). Exposure diversity as a design principle for recommender systems. *Information, Communication & Society*, 21(2), 191–207. <https://doi.org/10.1080/1369118X.2016.1271900>
- Helberger, N., Van Drunen, M., Eskens, S., & Bastian, M. (2020). A freedom of expression perspective on AI in the media – with a special focus on editorial decision making on social media platforms and in the news media. *European Journal of Law and Technology*, 11(3), 1–28. <https://ejlt.org/index.php/ejlt/article/view/752>
- Helberger, N., Van Drunen, M., Moeller, J., Vrijenhoek, S., & Eskens, S. (2022). Towards a normative perspective on journalistic AI: Embracing the messy reality of normative ideals. *Digital Journalism*, 10(10), 1605–1626. <https://doi.org/10.1080/21670811.2022.2152195>
- Helming, C. (2023, December 15). Microsoft's Bing chat: A source of misinformation on elections. *Algorithm Watch*. <https://algorithmwatch.org/en/microsofts-bing-source-misinformation-elections>
- Helsper, E. J. (2021). *The digital disconnect: The social causes and consequences of digital inequalities*. Sage. <https://doi.org/10.4135/9781526492982>
- Henderson, J., & Corry, M. (2020). Data literacy training and use for educational professionals. *Journal of Research in Innovative Teaching & Learning*, 14(2), 232–244. <https://doi.org/10.1108/JRIT-11-2019-0074>

- Heng, S., Tsilionis, K., Scharff, C., & Wautelet, Y. (2022). Understanding AI ecosystems in the Global South: The cases of Senegal and Cambodia. *International Journal of Information Management*, 64, 1–18. <https://doi.org/10.1016/j.ijinfomgr.2021.102454>
- Henningsen, A., & Krčál, A. (2024). *Public financing of news media in the EU: Final report*. Henningsen Consulting and Technopolis. <https://tinyurl.com/ms5xz73c>
- Hepp, A., Jarke, J., & Kramp, L. (2022). New perspectives in critical data studies: The ambivalences of data power – an introduction. In A. Hepp, J. Jarke, & L. Kramp (Eds.), *New perspectives in critical data studies: The ambivalences of data power* (pp. 1–23). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-96180-0_1
- Herman, E. S., & Chomsky, N. (1989). *Manufacturing consent: The political economy of the mass media*. Penguin.
- Hermida, A. (2023). *Journalism prepares for a post-search, post-social future*. Nieman Lab. <https://www.niemanlab.org/2023/12/journalism-prepares-for-a-post-search-post-social-future>
- Higgs, M. (2025, March 31). *Discrimination vs. neutrality: A case study in the middle of a trade war*. Tax Foundation. <https://taxfoundation.org/blog/digital-services-tax-europe-trade-war/>
- Hildebrandt, C., & Wiewiorra, L. (2024). The past, present, and future of (net) neutrality: A state of knowledge review and research agenda. *Journal of Information Technology*, 39(1), 167–193. <https://doi.org/10.1177/02683962231170891>
- Himma-Kadakas, M., & Ojames, I. (2022). Debunking false information: Investigating journalists’ fact-checking skills. *Digital Journalism*, 10(5), 866–887. <https://doi.org/10.1080/21670811.2022.2043173>
- Hinds, J., Williams, E. J., & Joinson, A. N. (2020). ‘It wouldn’t happen to me’: Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies*, 143, 1–55. <https://doi.org/10.1016/j.ijhcs.2020.102498>
- Hintz, A., Dencik, L., & Wahl-Jorgensen, K. (2019). *Digital citizenship in a datafied society*. Polity Press.
- Hirsch, M., Binder, A., & Matthes, J. (2025). A “drop in the ocean”? Emerging adults’ experiences and understanding of targeted political advertising on social media. *New Media & Society*, 1–26. Online First. <https://doi.org/10.1177/14614448241306455>
- Hlomani, H. (2023, September 12). *Future-proofing African democracy: Assessing electoral laws amidst AI advancements* [Blog post]. Research ICT Africa. <https://researchictafrica.net/2023/09/12/future-proofing-african-democracy-assessing-electoral-laws-amidst-ai-advancements>
- Hofmann, J. (2024). Desinformation als Symptom. Ein Überblick [Disinformation as a symptom: An overview]. In *Zentrum für Analyse/Forschung am Bundesamt für Verfassungsschutz (Hg.): Wissenschaftskonferenz 2023. Meinungsbildung 2.0 – Strategien im Ringen um Deutungshoheit im digitalen Zeitalter* [Center for Analysis/Research at the Federal Office for the Protection of the Constitution (ed.): Scientific Conference 2023. Opinion Formation 2.0 – Strategies in the Struggle for Interpretive Dominance in the Digital Age] (pp. 20–31). Bundesamt für Verfassungsschutz.
- Holt, J. (2023, November 9). *Information disorder in the Israel-Hamas war highlights shifts in the fight against online misinformation*. ISD Digital Dispatches. https://www.isdglobal.org/digital_dispatches/information-disorder-in-the-israel-hamas-war-highlights-shifts-in-the-fight-against-online-misinformation
- Hooper, L., Livingstone, S., & Pothong, K. (2022). *Problems with data governance in UK schools: The cases of Google classroom and class dojo*. Digital Futures Commission, 5Rights Foundation. <https://digitalfuturescommission.org.uk/wp-content/uploads/2022/08/Problems-with-data-governance-in-UK-schools.pdf>
- Hopkins, J. (2020). The concept of affordances in digital media. In H. Friese, M. Nolden, G. Rebane, & M. Schreier (Eds.), *Handbuch Soziale Praktiken und Digitale Alltagswelten* [Handbook of social practices and digital everyday worlds] (pp. 47–54). Springer. https://doi.org/10.1007/978-3-658-08460-8_67-1
- Horowitz, M., Milosavljević, M., & Van den Bulck, H. (2022). The use of artificial intelligence by public service media: Between advantages and threats. In C. El Morr (Ed.), *AI and society: Tensions and opportunities* (14 pp.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781003261247>

- Horowitz, M., Nieminen, H., Lehtisaari, K., & D'Arma, A. (Eds.). (2024). *Epistemic rights in the era of digital disruption*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-45976-4>
- Horst, H. A., Sargent, A., & Gaspard, I. (2024). Beyond extraction: Data strategies from the Global South. *New Media & Society*, 26(3), 1366–1383. <https://doi.org/10.1177/14614448231201651>
- Horten, M. (2023). Algorithms patrolling content: Where's the harm? *International Review of Law, Computers & Technology*, 38(1), 43–65. <https://doi.org/10.1080/13600869.2023.2221823>
- Horwitz, J. (2021). The Facebook files. *Wall Street Journal*. <https://www.wsj.com/articles/the-facebook-files-11631713039>
- Hoskins, G. T. (2024). Far right > digital rights: The precarity of free expression, internet access, net neutrality and data privacy in Bolsonaro's Brazil. *Javnost – The Public*, 31(2), 309–326. <https://doi.org/10.1080/13183222.2024.2346704>
- Houser, K., & Voss, W. G. (2018). GDPR: The end of Google and Facebook or a new paradigm in data privacy? *Richmond Journal of Law & Technology*, 25(1), 1–109. https://jolt.richmond.edu/files/2018/11/Houser_Voss-FE.pdf
- Howard, P. N., & Hussain, M. M. (2013). *Democracy's fourth wave? Digital media and the Arab Spring*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199936953.001.0001>
- Howard, P. N., Neudert, L.-M., Prakash, N., & Vosloo, S. (2021). *Digital misinformation / disinformation and children*. UNICEF Rapid Analysis Report. <https://www.unicef.org/globalinsight/media/2096/file/UNICEF-Global-Insight-Digital-Mis-Disinformation-and-Children-2021.pdf>
- HRW (Human Rights Watch). (2023). *Meta's broken promises: Systemic censorship of Palestine content on Instagram and Facebook*. <https://www.hrw.org/report/2023/12/21/metass-broken-promises/systemic-censorship-palestine-content-instagram-and>
- Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics*, 25(3), 493–516. <https://doi.org/10.1177/1940161219900126>
- Humprecht, E., Esser, F., Aelst, P. V., Staender, A., & Morosoli, S. (2023). The sharing of disinformation in cross-national comparison: Analyzing patterns of resilience. *Information, Communication & Society*, 26(7), 1342–1362. <https://doi.org/10.1080/1369118X.2021.2006744>
- Hung, T.-C., & Hung, T.-W. (2022). How China's cognitive warfare works: A frontline perspective of Taiwan's anti-disinformation wars. *Journal of Global Security Studies*, 7(4), 1–18. <https://doi.org/10.1093/jogss/ogac016>
- Hunsaker, A., & Hargittai, E. (2018). A review of Internet use among older adults. *New Media & Society*, 20(10), 3937–3954. <https://doi.org/10.1177/1461444818787348>
- Huszár, F., Ktena, S. I., O'Brien, C., Belli, L., Schlaikjer, A., & Hardt, M. (2022). Algorithmic amplification of politics on Twitter. *PNAS*, 119(1), 1–6. <https://www.pnas.org/doi/epdf/10.1073/pnas.2025334119>
- Hyzen, A. (2023). Propaganda and the Web 3.0: Truth and ideology in the digital age. *Nordic Journal of Media Studies*, 5(1), 49–67. <https://doi.org/10.2478/njms-2023-0004>
- Iamnitchi, A., Hall, L. O., Horawalavithana, S., Mubang, F., Ng, K. W., & Skvoretz, J. (2023). Modeling information diffusion in social media: Data-driven observations. *Frontiers in Big Data*, 6, 1–19. <https://doi.org/10.3389/fdata.2023.1135191>
- iea50. (2024). *Data centres and data transmission networks*. IEA (International Energy Agency). <https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks>
- IEEE (Institute of Electrical and Electronics Engineers). (2019). *Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems*. Version 2. <https://ieeexplore.ieee.org/document/9398613>
- Ihlebaek, K. A., & Sundet, V. S. (2023). Global platforms and asymmetrical power: Industry dynamics and opportunities for policy change. *New Media & Society*, 25(8), 2183–2200. <https://doi.org/10.1177/14614448211029662>
- Ihlebaek, K. A., Figenschou, T. U., Eldridge II, S. A., Frischlich, L., Cushion, S., & Holt, K. (2022). Understanding alternative news media and its contribution to diversity. *Digital Journalism*, 10(8), 1267–1282. <https://doi.org/10.1080/21670811.2022.2134165>

- Iliadis, A., & Acker, A. (2022). The seer and the seen: Surveying Palantir's surveillance platform. *The Information Society*, 38(5), 334–363. <https://doi.org/10.1080/01972243.2022.2100851>
- Innerarity, D. (2024). *Artificial intelligence and democracy*. UNESCO Regional Office, Montevideo and CLACSO (Latin American Council of Social Sciences). <https://unesdoc.unesco.org/ark:/48223/pf0000389736>
- Innes, H., Dawson, A., & Innes, M. (2023, October 10). *OSINT vs disinformation: The information threats 'arms race'*. Centre for Research and Evidence on Security Threats. <https://crestresearch.ac.uk/comment/osint-vs-disinformation-the-information-threats-arms-race>
- International IDEA (Institute for Democracy and Electoral Assistance). (2023). *The global state of democracy 2023: The new checks and balances*. <https://www.idea.int/gsod/2023>
- Introne, J., McKernan, B., Corsbie-Massay, C. L. P., Rohlinger, D., & Tripodi, F. B. (2024). Healthier information ecosystems: A definition and agenda. *Journal of the Association for Information Science and Technology*, 75(10), 1025–1040. <https://doi.org/10.1002/asi.24949>
- IPIE (International Panel on the Information Environment). (2023a). *Countermeasures for mitigating digital misinformation: A systemic review*. <https://www.ipie.info/reports-and-publications/reports-and-publications>
- IPIE (International Panel on the Information Environment). (2023b). *Platform responses to misinformation: A meta-analysis of data*. <https://www.ipie.info/reports-and-publications/reports-and-publications>
- IPIE (International Panel on the Information Environment). (2023c). *Trends in the global information environment*. <https://www.ipie.info/reports-and-publications/reports-and-publications>
- Ipsos, & UNESCO. (2023). *Survey on the impact of online disinformation and hate speech*. https://www.unesco.org/sites/default/files/medias/fichiers/2023/11/unesco_ipsos_survey_1.pdf
- Ipsos. (2025, June) *The IPSOS AI monitor 2025: A 30-country Ipsos global advisor survey*. <https://www.ipsos.com/sites/default/files/ct/publication/documents/2025-06/Ipsos-AI-Monitor-2025.pdf>
- Isin, E., & Ruppert, E. (2020). *Being digital citizens*. Rowman & Littlefield.
- IT for Change. (2023a). *Feminist perspectives on social media governance*. <https://itforchange.net/feminist-perspectives-on-social-media-governance-0>
- IT for Change. (2023b). *IT for change's submission to the forum on information and democracy's working group on AI and its implications for the information & communication space*. <https://tinyurl.com/46h5b22j>
- ITU (International Telecommunication Union). (2024a). *The ICT development index 2024: Measuring digital development*. https://www.itu.int/dms_pub/itu-d/opb/ind/D-IND-ICT_MDD-2024-3-PDF-E.pdf
- ITU (International Telecommunication Union). (2024b). *Statistics ITU*. <https://www.itu.int/en/ITU-D/Statistics/pages/stat/default.aspx>
- Ivanova, I. (2022, November 21). *These formerly banned Twitter accounts have been reinstated since Elon Musk took over*. CBS News. <https://www.cbsnews.com/news/twitter-accounts-reinstated-elon-musk-donald-trump-kanye-ye-jordan-peterson-kathy-griffin-andrew-tate>
- Jacob, N. (2012). The tendency to trust as individual predisposition – Exploring the associations between interpersonal trust, trust in the media and trust in institutions. *Communications*, 37(1), 99–120. <https://doi.org/10.1515/commun-2012-0005>
- Jackson, S. J., Liu, J., Singh, R., & Passi, S. (2025). Maintaining data infrastructures. In T. Venturini, A. Acker, J.-C. Plantin, & T. Walford (Eds.), *The Sage handbook of data and society* (pp. 19–35). Sage.
- Jagannatha, A., Rawat, B. P. S., & Yu, H. (2021). Membership inference attack susceptibility of clinical language models. *ArXiv*. <https://doi.org/10.48550/arXiv.2104.08305>
- Jain, A. (2023, May 13). *We wrote to the election commissioner of India on the use of facial recognition technology for voter verification in Karnataka*. Internet Freedom Foundation. <https://internetfreedom.in/we-wrote-to-the-election-commissioner-of-india-on-the-use-of-facial-recognition-technology-for-voter-verification>
- Jakobsson, P., & Stiernstedt, F. (2023). Trust and the media: Arguments for the (ir)relevance of a concept. *Journalism Studies*, 24(4), 479–495. <https://doi.org/10.1080/1461670X.2023.2169191>

- James, D. (2022). *Rigging the rules: How big tech uses stealth “trade” agreements and how we can stop them*. State of Big Tech.
<https://projects.itforchange.net/state-of-big-tech/rigging-the-rules-how-big-tech-uses-stealth-trade-agreements-to-undermine-and-prevent-digitalization-in-public-interest-and-how-we-can-stop-them>
- Jamieson, K. H., & Albarracin, D. (2020). The relation between media consumption and misinformation at the outset of the SARS-CoV-2 pandemic in the US. *The Harvard Kennedy School (HKS) Misinformation Review*, 1(2), 1–23.
<https://nrs.harvard.edu/urn-3:HUL.InstRepos:42661740>
- Jamieson, K. H., & Cappella, J. N. (2008). *Echo chamber: Rush Limbaugh and the conservative media establishment*. Oxford University Press.
- Jandrić, P. (2019). The postdigital challenge of critical media literacy. *The International Journal of Critical Media Literacy*, 1(1), 26–37. <https://doi.org/10.1163/25900110-00101002>
- Jasanoff, S. (2015). Future imperfect: Science, technology, and the imaginations of modernity. In S. Jasanoff, & S.-H. Kim (Eds.), *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power* (pp. 1–34). The University of Chicago Press.
- Jauniaux, L., & Lebourges, M. (2019). Zero rating and end-users’ freedom of choice: An economic analysis. *Digital Policy, Regulation and Governance*, 21(2), 115–128.
<https://doi.org/10.1108/DPRG-06-2018-0030>
- Jayasinghe, G., Bandaranayake, R., Galpaya, H., Ababakirov, A., & De Silva, R. (2022). Asia. In H. Wasserman (Ed.), *Meeting the challenges of information disorder in the Global South* (pp. 190–245). IDRC Research Results. <https://idl-bnc-idrc.dspacedirect.org/bitstreams/aafc0871-9bfe-4d11-bd8e-eb7921064621/download>
- Jeong, S.-H., Cho, H., & Hwang, Y. (2012). Media literacy interventions: A meta-analytic review. *Journal of Communication*, 62(3), 454–472.
<https://doi.org/10.1111/j.1460-2466.2012.01643.x>
- Jeppesen, S. (2016). Understanding alternative media power: Mapping content & practice to theory, ideology, and political action. *Democratic Communiqué*, 27(1), 54–77.
<https://journals.flvc.org/demcom/article/view/88322>
- Jeppesen, S., Hoechsmann, M., ulthiin, i. h., VanDyke, D., & McKee, M. (2022). *The capitol riots: Digital media, disinformation, and democracy under attack*. Routledge.
<https://doi.org/10.4324/9781003246862>
- Jhaver, S., & Zhang, A. X. (2023). Do users want platform moderation or individual control? Examining the role of third-person effects and free speech support in shaping moderation preferences. *New Media & Society*, 27(5), 2930–2950.
<https://doi.org/10.1177/14614448231217993>
- Jiang, B., Tan, Z., Nirmal, A., & Liu, H. (2024). Disinformation detection: An evolving challenge in the age of LLMs. *Proceedings of the 2024 SIAM international conference on data mining (SDM)* (pp. 427–435). Society for Industrial and Applied Mathematics.
<https://epubs.siam.org/doi/pdf/10.1137/1.9781611978032.50>
- Jiang, M., & Belli, L. (Eds.). (2025). *Digital sovereignty in the BRICS countries: How the Global South and emerging power alliances are reshaping digital governance*. Cambridge University Press. <https://doi.org/10.1017/9781009531085>
- Jin, Y., & Austin, L. (Eds.). (2022). *Social media and crisis communication*. Routledge.
<https://doi.org/10.4324/9781003043409>
- Jindal, T., & Nigam, A. (2020). *Data stewardship for non-personal data in India: A position paper on data trusts*. Vidhi Centre for Legal Policy, Indraprastha Institute of Information Technology, Cyril Amarch and Mangaldas and Omidyar Network India.
<https://vidhilegalpolicy.in/research/data-stewardship-for-non-personal-data-in-india/>
- Johnson, G. M. (2023). Are algorithms value-free? Feminist theoretical virtues in machine learning. *Journal of Moral Philosophy*, 21(1-2), 27–61.
<https://doi.org/10.1163/17455243-20234372>
- Jones-Jang, S. M., & Chung, M. (2024). Can we blame social media for polarization? Counter-evidence against filter bubble claims during the COVID-19 pandemic. *New Media & Society*, 26(6), 3370–3389. <https://doi.org/10.1177/14614448221099591>
- Jones-Jang, S. M., Mortensen, T., & Liu, J. (2021). Does media literacy help identification of fake news? Information literacy helps, but other literacies don’t. *American Behavioral Scientist*, 65(2), 371–388. <https://doi.org/10.1177/0002764219869406>

- Jones, M. O. [@marcowenjones]. (2024, August 9). *New analysis showing ALL tweets mentioning both “Muslim” & “UK” with over 500 retweets since 29 Jul.* [Post]. X.
<https://x.com/marcowenjones/status/1821881958537871568>
- Jordan, S. (2017). Evaluating zero-rating and associated throttling practices under the open internet order. *Journal of Information Policy*, 7, 450–507.
<https://doi.org/10.5325/jinfolpoli.7.2017.0450>
- Jørgensen, R. F. (2017). What platforms mean when they talk about human rights. *Policy & Internet*, 9(3), 280–296. <https://doi.org/10.1002/poi3.152>
- Joris, G., Grove, F. D., Van Damme, K., & De Marez, L. (2021). Appreciating news algorithms: Examining audiences’ perceptions to different news selection mechanisms. *Digital Journalism*, 9(5), 589–618. <https://doi.org/10.1080/21670811.2021.1912626>
- Jung, E. H., & Sundar, S. S. (2016). Senior citizens on Facebook: How do they interact and why? *Computers in Human Behavior*, 61, 27–35. <https://doi.org/10.1016/j.chb.2016.02.080>
- Jungherr, A., & Schroeder, R. (2021). Disinformation and the structural transformations of the public arena: Addressing the actual challenges to democracy. *Social Media + Society*, 7(1), 1–13. <https://doi.org/10.1177/2056305121988928>
- Jungherr, A., & Schroeder, R. (2023). Artificial intelligence and the public arena. *Communication Theory*, 33(2-3), 164–173. <https://doi.org/10.1093/ct/qtad006>
- Just, N. (2018). Governing online platforms: Competition policy in times of platformization. *Telecommunications Policy*, 42(5), 386–394. <https://doi.org/10.1016/j.telpol.2018.02.006>
- Just, N. (2022). Which is to be master? Competition law or regulation in platform markets. *International Journal of Communication*, 16, 504–524.
<https://ijoc.org/index.php/ijoc/article/view/13095/3657>
- Privacy Law Library. (2017). *Justice K.S. Puttaswamy (Retd.) & Anr. vs. Union of India & Ors* (Judge Bench Civil Writ Petition, Case brief). <https://tinyurl.com/26kpi7zu>
- Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment.* William Collins.
- Kakkar, J. M. (2023). Tackling misinformation in emerging economies and the Global South: Exploring approaches for the Indian context. In F. Fukuyama, & M. Schaake (Eds.), *Digital technologies in emerging countries* (pp. 48–62). Cyber Policy Centre, Freeman Spogli Institute; Stanford Law School, Stanford University. <https://csgdelhi.s3.ap-south-1.amazonaws.com/uploads/digital-technologies-in-emerging-countries-digital-457.pdf>
- Kalathil, S., & Boas, T. C. (2003). *Open networks, closed regimes: The impact of the internet on authoritarian rule.* Carnegie Endowment for International Peace.
- Kamp, C. (2023). *Joint statement of the informal OSCE Group of friends on safety of journalists.* US Mission to the OSCE. <https://osce.usmission.gov/joint-statement-of-the-informal-osce-group-of-friends-on-safety-of-journalists>
- Kantar Media. (2016). *Brand and trust in a fragmented news environment.* Reuters Institute for the Study of Journalism, University of Oxford.
<https://reutersinstitute.politics.ox.ac.uk/our-research/brand-and-trust-fragmented-news-environment>
- Kapantai, E., Christopoulou, A., Berberidis, C., & Peristeras, V. (2021). A systematic literature review on disinformation: Toward a unified taxonomical framework. *New Media & Society*, 23(5), 1301–1326. <https://doi.org/10.1177/1461444820959296>
- Kaplan, J. (2025, January 7). *More speech and fewer mistakes.* Meta.
<https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/>
- Karlsen, R., Beyer, A., & Steen-Johnsen, K. (2020). Do high-choice media environments facilitate news avoidance? A longitudinal study 1997–2016. *Journal of Broadcasting & Electronic Media*, 64(5), 794–814. <https://doi.org/10.1080/08838151.2020.1835428>
- Karpf, D. (2020). How digital disinformation turned dangerous. In S. Livingston, & W. L. Bennett (Eds.), *The disinformation age* (pp. 153–168). Cambridge University Press.
<https://doi.org/10.1017/9781108914628>
- Kaskazi, A., & Kitzie, V. (2023). Engagement at the margins: Investigating how marginalized teens use digital media for political participation. *New Media & Society*, 25(1), 72–94.
<https://doi.org/10.1177/14614448211009460>

- Katsirea, I. (2018). "Fake news": Reconsidering the value of untruthful expression in the face of regulatory uncertainty. *Journal of Media Law*, 10(2), 159–188.
<https://doi.org/10.1080/17577632.2019.1573569>
- Katz, E. (1957). The two-step flow of communication: An up-to-date report on an hypothesis. *Public Opinion Quarterly*, 21(1), 61–78. <https://doi.org/10.1086/266687>
- Katzenbach, C. (2021). "AI will fix this" – The technical, discursive, and political turn to AI in governing communication. *Big Data & Society*, 8(2), 1–8.
<https://doi.org/10.1177/20539517211046182>
- Kaye, D. (2015, May). *Report of the special rapporteur on the promotion and protection of the right to freedom of opinion and expression*. UN Human Rights Council.
<https://www.ohchr.org/en/special-procedures/sr-freedom-of-opinion-and-expression>
- Kaye, D. (2018). *Mandate of the special rapporteur on the promotion and protection of the right to freedom of opinion and expression: Letter by David Kay*. OHCHR Letter OL-ITA-1-2018. <https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/Legislation/OL-ITA-1-2018.pdf>
- Keller, T. R., & Klinger, U. (2019). Social bots in election campaigns: Theoretical, empirical, and methodological implications. *Political Communication*, 36(1), 171–189.
<https://doi.org/10.1080/10584609.2018.1526238>
- Kennedy, I., Wack, M., Beers, A., Schafer, J. S., Garcia-Camargo, I., Spiro, E. S., & Starbird, K. (2022). Repeat spreaders and election delegitimization: A comprehensive dataset of misinformation tweets from the 2020 US election. *Journal of Quantitative Description: Digital Media*, 2, 1–49. <https://doi.org/10.51685/jqd.2022.013>
- Kennis, A. (2024). Elite news content (still) matters now more than ever before: Social movements, social media, elite news media, and the media dependence model. In J. Pedro-Carañana, R. Gómez, T. F. Corrigan, & F. R. Sierra Caballero (Eds.), *Political economy of media and communication: Methodological approaches* (pp. 224–244). Routledge.
<https://doi.org/10.4324/9781003385516>
- Kerber, W. (2020). *From (horizontal and sectoral) data access solutions towards data governance systems* [Joint Discussion Paper No. 40-2020. Series in Economics. Universities of Aachen, Gießen, Göttingen Kassel, Marburg, Siegen].
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3681263
- Keremoglu, E., & Weidmann, N. B. (2020). How dictators control the internet: A review essay. *Comparative Political Studies*, 53(10-11), 1690–1703.
<https://doi.org/10.1177/0010414020912278>
- Kertysova, K. (2018). Artificial intelligence and disinformation: How AI changes the way disinformation is produced, disseminated, and can be countered. *Security and Human Rights*, 29(1-4), 55–81. <https://doi.org/10.1163/18750230-02901005>
- Kettemann, M. C. (2022). *The normative order of the internet: A theory of rule and regulation online*. Oxford University Press. <https://doi.org/10.1093/oso/9780198865995.001.0001>
- Kettemann, M. C., & Lachmayer, K. (Eds.). (2021). *Pandemocracy in Europe: Power, parliaments and people in times of COVID-19*. Bloomsbury.
- Kettemann, M. C., & Schulz, W. (2023). *Platform://Democracy – Perspectives on platform power, public values and the potential of social media councils*. Verlag Hans-Bredow-Institut.
<https://graphite.page/platform-democracy-report/>
- Kim, K., & Moon, S.-I. (2021). When algorithmic transparency failed: Controversies over algorithm-driven content curation in the South Korean digital environment. *American Behavioral Scientist*, 65(6), 847–862. <https://doi.org/10.1177/0002764221989783>
- Kim, S. J. (2016). A repertoire approach to cross-platform media use behavior. *New Media & Society*, 18(3), 353–372. <https://doi.org/10.1177/1461444814543162>
- Kinzelbach, K., Lindberg, S. I., & Lott, L. (2024). *Academic freedom index update 2024*. Friedrich-Alexander-Universität; V-Dem Institute.
https://academic-freedom-index.net/research/Academic_Freedom_Index_Update_2024.pdf
- Kira, B., Sinha, V., & Srinivasan, S. (2021). Regulating digital ecosystems: Bridging the gap between competition policy and data protection. *Industrial and Corporate Change*, 30(5), 1337–1360. <https://doi.org/10.1093/icc/dtab053>
- Klapper, J. T. (1960). *The effects of mass communication*. Free Press.
- Klein, N. (2008). *The shock doctrine: The rise of disaster capitalism*. Picador.

- Klein, N. (2020, May 13). Naomi Klein: How big tech plans to profit from the pandemic. *The Guardian*. <https://www.theguardian.com/news/2020/may/13/naomi-klein-how-big-tech-plans-to-profit-from-coronavirus-pandemic>
- Klonick, K. (2017). The new governors: The people, rules, and processes governing online speech. *Harvard Law Review*, 131(6), 1598–1670. https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf
- Knight, W. (2024, June 4). OpenAI employees warn of a culture of risk and retaliation. *Wired*. <https://www.wired.com/story/openai-right-to-warn-open-letter-ai-risk>
- Knuutila, A., Neudert, L.-M., & Howard, P. N. (2022). Who is afraid of fake news? Modeling risk perceptions of misinformation in 142 countries. *Harvard Kennedy School (HKS) Misinformation Review*, 3(3), 1–13. <https://doi.org/10.37016/mr-2020-97>
- Koene, A., Smith, A. L., Egawa, T., Mandalh, S., & Hatada, Y. (2018). IEEE P70xx, establishing standards for ethical technology. *Proceedings of KDD*. https://www.kdd.org/kdd2018/files/project-showcase/KDD18_paper_1743.pdf
- Kohring, M., & Matthes, J. (2007). Trust in news media: Development and validation of a multidimensional scale. *Communication Research*, 34(2), 231–252. <https://doi.org/10.1177/0093650206298071>
- Kokas, A. (2022). Becoming a cyber sovereign: China's politics of data governance. In A. Kokas, *Trafficking data: How China is winning the battle for digital sovereignty* (pp. 51–C53. P72). Oxford Academic.
- Kokshagina, O., Reinecke, P. C., & Karanasios, S. (2023). To regulate or not to regulate: Unravelling institutional tussles around the regulation of algorithmic control of digital platforms. *Journal of Information Technology*, 38(2), 160–179. <https://doi.org/10.1177/02683962221114408>
- Kop, M. (2020). *Shaping the law of AI: Transatlantic perspectives*. Stanford Law School. <https://law.stanford.edu/publications/no-65-shaping-the-law-of-ai-transatlantic-perspectives>
- Kostygina, G., Kim, Y., Seeskin, Z., LeClere, F., & Emery, S. (2023). Disclosure standards for social media and generative artificial intelligence research: Toward transparency and replicability. *Social Media + Society*, 9(4), 1–12. <https://doi.org/10.1177/20563051231216947>
- Kothari, A., & Cruikshank, S. A. (2022). Artificial intelligence and journalism: An agenda for journalism research in Africa. *African Journalism Studies*, 43(1), 17–33. <https://doi.org/10.1080/23743670.2021.1999840>
- Kozyreva, A., Lorenz-Spreen, P., Herzog, S. M., Ecker, U. K. H., Lewandowsky, S., & Hertwig, R., Ali, A., Bak-Coleman, J., Barzilai, S., Basol, M., Berinsky, A. J., Betsch, C., Cook, J., Fazio, L. K., Geers, M., Guess, A. M., Huang, H., Larreguy, H., Maertens, R., ... Wineberg, S. (2024). Toolbox of individual-level interventions against online misinformation. *Nature Human Behavior*, 8(6), 1044–1052. <https://doi.org/10.1038/s41562-024-01881-0>
- Krämer, J., & Peitz, M. (2018). A fresh look at zero-rating. *Telecommunications Policy*, 42(7), 501–513. <https://doi.org/10.1016/j.telpol.2018.06.005>
- Kreiss, D., & McGregor, S. C. (2023). A review and provocation: On polarization and platforms. *New Media & Society*, 26(1), 556–579. <https://doi.org/10.1177/14614448231161880>
- Kreps, S., McCain, R. M., & Brundage, M. (2022). All the news that's fit to fabricate: AI-generated text as a tool of media misinformation. *Journal of Experimental Political Science*, 9(1), 104–117. <https://doi.org/10.1017/XPS.2020.37>
- Kretschmer, M., Meletti, B., Bently, L., Cirfodelli, G., Eben, M., Erickson, K., Iramina, A., Li, Z., McDonagh, L., Perot, E., Porangaba, L., & Thomas, A. (2025). *Copyright and AI: Response by CREATE Centre to the UK Government's Consultation [CREATE Working Paper 2025/2]*. <https://zenodo.org/records/14931964>
- Kristensen, L. M., & Hartley, J. M. (2023). The infrastructure of news: Negotiating infrastructural capture and autonomy in data-driven news distribution. *Media and Communication*, 11(2), 307–318. <https://doi.org/10.17645/mac.v11i2.6388>
- Kroet, C. (2024, August 6). *Meta urged to answer EU questions on CrowdTangle*. Euronews. <https://www.euronews.com/next/2024/08/16/meta-urged-to-answer-eu-questions-on-crowdtangle>
- Krotoszynski, Jr., Ronald J., Koltay, A., & Garden, C. (Eds.). (2025). *Disinformation, misinformation, and democracy: Legal approaches in comparative context*. Cambridge University Press. <https://doi.org/10.1017/9781009373272>

- Krouwel, A., Kutysiński, Y., van Prooijen, J.-W., Martinsson, J., & Markstedt, E. (2017). Does extreme political ideology predict conspiracy beliefs, economic evaluations and political trust? Evidence From Sweden. *Journal of Social and Political Psychology*, 5(2), 435–462. <https://doi.org/10.5964/jssp.v5i2.745>
- Księżak, P., & Wojtczak, S. (2023). *Toward a conceptual network for the private law of artificial intelligence* (Vol. 51). Springer. <https://doi.org/10.1007/978-3-031-19447-4>
- Kulesz, M. (2018). *Culture, platforms and machines: The impact of artificial intelligence on the diversity of cultural expressions* [UNESCO Report to Intergovernmental Committee for the Protection and Promotion of the Diversity of Cultural Expressions, DCE/18/12.IGC/INF.4]. <https://www.unesco.org/creativity/en>
- Kumar, A., & Thussu, D. (2023). Media, digital sovereignty and geopolitics: The case of the TikTok ban in India. *Media, Culture & Society*, 45(8), 1583–1599. <https://doi.org/10.1177/01634437231174351>
- Kumar, R., Desai, V. S., & Koshy, N. S. (2023). *Creating sustainable data cooperatives in the Global South: Frameworks for institutional support* [Research brief]. IT for Change, India. <https://tinyurl.com/43mxkhsu>
- Kuznetsova, E., & Makhortykh, M. (2023). Blame it on the algorithm? Russian government-sponsored media and algorithmic curation of political information on Facebook. *International Journal of Communication*, 17, 971–992. <https://ijoc.org/index.php/ijoc/article/view/18687>
- Kyriakidou, M., Cushion, S., Hughes, C., & Morani, M. (2023). Questioning fact-checking in the fight against disinformation: An audience perspective. *Journalism Practice*, 17(10), 2123–2139. <https://doi.org/10.1080/17512786.2022.2097118>
- la Cour, C. (2020). Theorising digital disinformation in international relations. *International Politics*, 57(4), 704–723. <https://doi.org/10.1057/s41311-020-00215-x>
- Labio-Bernal, A., & Manzano-Zambruno, L. (2023). 8M demonstrations, the Spanish far right and the pandemic in a hybrid media system. *Media and Communication*, 11(1), 114–124. <https://idus.us.es/handle/11441/142909>
- Labio-Bernal, A., & Romero-Domínguez, L. R. (2022). *Periodismo SOS – Del filantropocapitalismo al poder que mata* [SOS Journalism – From philanthrocapitalism to the power that kills]. Ediciones Pirámide.
- LaForge, G., & Gruver, P. (2023, October). *Governing the digital future*. Planetary Politics. <https://core.ac.uk/download/588303818.pdf>
- Lamdan, S. (2022). *Data cartels: The Companies that control and monopolize our information*. Stanford University Press.
- Landau, M. (2020). The astounding growth of “big tech” and the lack of enforcement of the intellectual property, antitrust and contract laws. *Albany Law Journal of Science and Technology*, 30(3), 1–61. <https://heinonline.org/HOL/LandingPage?handle=hein.journals/albnyst30&div=15&id=&page=>
- Lane, J., Stodden, V., Bender, S., & Nissenbaum, H. (Eds.). (2014). *Privacy, big data and the public good: Frameworks for engagement*. Cambridge University Press. <https://doi.org/10.1017/CBO9781107590205>
- Langer, A. I., & Gruber, J. B. (2021). Political agenda setting in the hybrid media system: Why legacy media still matter a great deal. *The International Journal of Press/Politics*, 26(2), 313–340. <https://doi.org/10.1177/1940161220925023>
- Lanier, J., & Weyl, E. G. (2018, September 26). A blueprint for a better digital society. *Harvard Business Review*, 2–19. <https://hbr.org/2018/09/a-blueprint-for-a-better-digital-society>
- Lanzing, M. (2023). Traveling technology and perverted logics: Conceptualizing Palantir’s expansion into health as sphere transgression. *Information, Communication and Society*, 27(15), 2617–2633. <https://doi.org/10.1080/1369118x.2023.2279557>
- Larivière, V., Haustein, S., & Mongeon, P. (2015). The oligopoly of academic publishers in the digital era. *PLOS ONE*, 10(6), 1–15. <https://doi.org/10.1371/journal.pone.0127502>
- Larsen, B. C., & Küspert, S. (2024, May 21). *Regulating general-purpose AI: Areas of convergence and divergence across the EU and the US*. Brookings Institution Research. <https://www.brookings.edu/articles/regulating-general-purpose-ai-areas-of-convergence-and-divergence-across-the-eu-and-the-us>
- Lasswell, H. D. (1971). *Propaganda technique in World War I*. MIT Press.

- Lauer, L., & Graves, L. (2024). How to grow a transnational field: A network analysis of the global fact-checking movement. *New Media & Society*, 27(6), 3505–3523. <https://doi.org/10.1177/14614448241227856>
- Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. *Computers and Education: Artificial Intelligence*, 3, 1–15. <https://doi.org/10.1016/j.caeai.2022.100101>
- Lawder, D. (2023, November 8). *US suspends Indo-Pacific talks on key aspects of digital trade –lawmakers*. Reuters. www.reuters.com/business/finance/us-suspends-indo-pacific-talks-key-aspects-digital-trade-lawmakers-2023-11-08
- Lawrence, S. E.-L. (2022, November 4). *The Majority World – what’s in a phrase?* [Blog post]. Philanthropy Australia. <https://www.philanthropy.org.au/news-and-stories/the-majority-world-whats-in-a-phrase>
- Lecheler, S., & de Vreese, C. H. (2017). News media, knowledge, and political interest: Evidence of a dual role from a field experiment: News media, knowledge, and political interest. *Journal of Communication*, 67(4), 545–564. <https://doi.org/10.1111/jcom.12314>
- Ledford, H. (2019). Millions of black people affected by racial bias in health-care algorithms. *Nature*, 574(7780), 608–609. <https://doi.org/10.1038/d41586-019-03228-6>
- Lee, D., & Shin, D.-H. (2016). The effects of network neutrality on the incentive to discriminate, invest, and innovate: A literature review. *info*, 18(3), 42–57. <https://doi.org/10.1108/info-12-2015-0053>
- Lee, J. (2024). CCPA/CPRA: Consumers bear the burden as companies bear the crown. *Hastings International & Comparative Law Review*, 47(2), 129–152. https://repository.uclawsf.edu/hastings_international_comparative_law_review/vol47/iss2/5
- Lee, S., & Valenzuela, S. (2024). A self-righteous, not a virtuous, circle: Proposing a new framework for studying media effects on knowledge and political participation in a social media environment. *Social Media + Society*, 10(2), 1–7. <https://doi.org/10.1177/20563051241257953>
- Lehdonvirta, V. (2022). *Cloud empires: How digital platforms are overtaking the state and how we can regain control*. MIT Press. <https://doi.org/10.7551/mitpress/14219.001.0001>
- Lehuedé, S. (2022). Territories of data: Ontological divergences in the growth of data infrastructure. *Tapuya: Latin American Science, Technology and Society*, 5(1), 1–18. <https://doi.org/10.1080/25729861.2022.2035936>
- Lemke, T., & Habegger, M. W. (2022). Foreign interference and social media networks: A relational approach to studying contemporary Russian disinformation. *Journal of Global Security Studies*, 7(2), 1–25. <https://doi.org/10.1093/jogss/ogac004>
- Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes. *Philosophy & Technology*, 31(4), 611–627. <https://doi.org/10.1007/s13347-017-0279-x>
- Lesh, M. (2023). *Breaking the news? Should digital platforms be required to fund news publishers?* [IEA Discussion paper No. 119]. Institute of Economic Affairs. https://iea.org.uk/wp-content/uploads/2023/02/DP119_Breaking-the-news_web-2.pdf
- Leslie, D., Katell, M., Aitken, M., Singh, J., Briggs, M., Powell, R., Rincon, C., Perini, A., & Jayadeva, S. (2022). *Data justice in practice: A guide for impacted communities*. Global Partnership on Artificial Intelligence. <https://dx.doi.org/10.2139/ssrn.4080046>
- Lewandowsky, S., Armaos, K., Bruns, H., Schmid, P., Holford, D. L., Hahn, U., Al-Rawi, A., Sah, S., & Cook, J. (2022). When science becomes embroiled in conflict: Recognizing the public’s need for debate while combating conspiracies and misinformation. *The ANNALS of the American Academy of Political and Social Science*, 700(1), 26–40. <https://doi.org/10.1177/00027162221084663>
- Lewkowiz, G., & Sarf, R. (2024). Taking technical standardization of fundamental rights seriously for trustworthy artificial intelligence. *La Revue des Juristes de Sciences Po*, 25, 42–46. <https://centrepereleman.be/content/uploads/2024/03/GL-RS-AI-Technical-standards-and-fundamental-rights.pdf>
- LexisNexis. (2024). *Sources of local and international news and business information*. Retrieved July 8, 2025, from <https://www.lexisnexis.ca/en-ca/products/nexis.page>
- Li, W. Y. (2023). Regulatory capture’s third face of power. *Socio-Economic Review*, 21(2), 1217–1245. <https://doi.org/10.1093/ser/mwad002>

- Lim, G., & Bradshaw, S. (2023). *Chilling legislation: Tracking the impact of “fake news” laws on press freedom internationally*. Center for International Media Assistance.
<https://www.cima.ned.org/publication/chilling-legislation/>
- Lin, B., & Lewis, S. C. (2022). The one thing journalistic AI just might do for democracy. *Digital Journalism*, 10(10), 1627–1649. <https://doi.org/10.1080/21670811.2022.2084131>
- Linden, T., Khandelwal, R., Harkous, H., & Fawaz, K. (2020). The privacy policy landscape after the GDPR. *ArXiv*. <https://doi.org/10.48550/arXiv.1809.083696>
- Linville, D. L., & Warren, P. L. (2019). *Touched by the trolls: How and why a coordinated information operation interacts with outsiders*. <https://tinyurl.com/2vyx9r2j>
- Linville, D. L., & Warren, P. L. (2020). Troll factories: Manufacturing specialized disinformation on Twitter. *Political Communication*, 37(4), 447–467.
<https://doi.org/10.1080/10584609.2020.1718257>
- Liu, Y., Zhang, K., Li, Y., Yan, Z., Gao, C., Chen, R., Yuan, Z., Huang, Y., Sun, H., Jianfeng Gao, Lifang He, Lichao Sun (2024). Sora: A review on background, technology, limitations, and opportunities of large vision models. *ArXiv*. <https://doi.org/10.48550/arXiv.2402.17177>
- Liveriero, F. (2020). Epistemic injustice in the political domain: Powerless citizens and institutional reform. *Ethical Theory and Moral Practice*, 23(5), 797–813.
<https://doi.org/10.1007/s10677-020-10097-w>
- Livingstone, S. (1996). On the continuing problems of media effects research. In J. Curran, & M. Gurevitch (Eds.), *Mass media and society* (pp. 305–324). Edward Arnold.
- Livingstone, S. (1998). Relationships between media and audiences: Prospects for audience reception studies. In T. Liebes, & J. Curran (Eds.), *Media, ritual and identity: Essays in honor of Elihu Katz* (pp. 237–255). Routledge.
- Livingstone, S. (2008). Engaging with media – a matter of literacy? *Communication, Culture & Critique*, 1(1), 51–62. <https://doi.org/10.1111/j.1753-9137.2007.00006.x>
- Livingstone, S. (2023). *What’s the best we can expect of media literacy? From protectionism to human rights and flourishing* [Blog post]. Media Literacy.
<https://blogs.lse.ac.uk/medialse/2023/12/13/whats-the-best-we-can-expect-of-media-literacy-from-protectionism-to-human-rights-and-flourishing>
- Livingstone, S., & Blum-Ross, A. (2020). *Parenting for a digital future: How hopes and fears about technology shape children’s lives*. Oxford University Press.
<https://doi.org/10.1093/oso/9780190874698.001.0001>
- Livingstone, S., Cantwell, N., Özkul, D., Shekhawat, G., & Kidron, B. (2024). *The best interests of the child in the digital environment*. Digital Futures for Children Centre, London School of Economics; Political Science and 5Rights Foundation.
<https://5rightsfoundation.com/uploads/dfc-report-best-interests-of-the-child.pdf>
- Livingstone, S., Mascheroni, G., & Stoilova, M. (2023a). The outcomes of gaining digital skills for young people’s lives and wellbeing: A systematic evidence review. *New Media & Society*, 25(5), 1176–1202. <https://doi.org/10.1177/14614448211043189>
- Livingstone, S., Ólafsson, K., & Pothong, K. (2023b). Digital play on children’s terms: A child rights approach to designing digital experiences. *New Media & Society*, 27(3), 1465–1485.
<https://doi.org/10.1177/14614448231196579>
- Livio, O., & Cohen, J. (2018). ‘Fool me once, shame on you’: Direct personal experience and media trust. *Journalism*, 19(5), 684–698. <https://doi.org/10.1177/1464884916671331>
- Llansó, E. J., van Hoboken, J., & Harambam, J. (2020). *Artificial intelligence, content moderation, and freedom of expression*. TransAtlantic Working Group on Content Moderation Online and Freedom of Expression, Annenberg Public Policy Center and Institute for Information Law, University of Amsterdam. <https://tinyurl.com/5ecchvnx>
- Local Progress. (2024, January 9). *Creating traffic safety: A policy memo for local elected leaders*.
<https://localprogress.org/resource/creating-traffic-safety-a-policy-memo-for-local-elected-leaders>
- Lockie, S. (2017). Post-truth politics and the social sciences. *Environmental Sociology*, 3(1), 1–5. <https://doi.org/10.1080/23251042.2016.1273444>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *CHI’20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376727>

- Longo, L., Brcic, M., Cabitza, F., Choi, J., Confalonieri, R., Del Ser, J., Guidotti, R., Hayashi, Y., Herrera, F., Holzinger, A., Jiang, R., Khosravi, H., Lecue, F., Malgieri, G., Páez, A., Samek, W., Schneider, J., Speith, T., & Stumpf, S. (2024). Explainable artificial intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions. *ArXiv*. <https://doi.org/10.1016/j.inffus.2024.102301>
- Loos, E., & Ivan, L. (2023). Using media literacy to fight digital fake news in later life: A mission impossible? In Q. Gao & J. Zhou (Eds.), *Human aspects of IT for the aged population. HCII 2023. Lecture Notes in Computer Science*, vol 14042. Springer. https://doi.org/10.1007/978-3-031-34866-2_18
- LSE Truth, Trust & Technology Commission. (2018). *Tackling the information crisis: A policy framework for media system resilience*. LSE Commission on Truth, Trust and Technology. <https://www.lse.ac.uk/media-and-communications/assets/documents/research/T3-Report-Tackling-the-Information-Crisis-v6.pdf>
- Luchs, I. (2023). AI for all? Challenging the democratization of machine learning. *Minor Tech: A Peer-Reviewed Journal About AI*, 12(1), 135–145. <https://doi.org/10.7146/aprja.v12i1.140445>
- Lunga, C. M., & Mthembu, M. V. (2019). Investigating the source and strategies adopted by mainstream media in combating fake news in the Kingdom of Eswatini. *African Journalism Studies*, 40(4), 96–111. <https://doi.org/10.1080/23743670.2019.1664606>
- Mabweazara, H. M., & Pearson, B. (Eds.). (2024). *Media capture in Africa and Latin America: Power and resistance*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-68962-8>
- Mabweazara, H. M., Muneri, C. T., & Ndlovu, F. (2020). News “media capture”, relations of patronage and clientelist practices in sub-Saharan Africa: An interpretive qualitative analysis. *Journalism Studies*, 21(15), 2154–2175. <https://doi.org/10.1080/1461670X.2020.1816489>
- MacCarthy, M. (2023). *Regulating digital industries: How public oversight can encourage competition, protect privacy, and ensure free speech*. Brookings Institution Press.
- Macdonald, S., Correia, S. G., & Watkin, A.-L. (2019). Regulating terrorist content on social media: Automation and the rule of law. *International Journal of Law in Context*, 15(2), 183–197. <https://doi.org/10.1017/S1744552319000119>
- Machill, S. A. (2020). *Biased artificial intelligence: Algorithmic fairness and human perception of biased AI*. NOVA Information Management School, Universidade Nova de Lisboa. <https://run.unl.pt/bitstream/10362/109738/1/TGI0357.pdf>
- MacKenzie, D., Caliskan, K., & Rommerskirchen, C. (2023). The longest second: Header bidding and the material politics of online advertising. *Economy and Society*, 52(3), 554–578. <https://doi.org/10.1080/03085147.2023.2238463>
- Madan, G. D. (2021). Understanding misinformation in India: The case for a meaningful regulatory approach for social media platforms. *ArXiv*. <https://doi.org/10.48550/arXiv.2207.01508>
- Madrid-Morales, D., & Wasserman, H. (2022). Research methods in comparative disinformation studies. In H. Wasserman, & D. Madrid-Morales (Eds.), *Disinformation in the Global South* (pp. 41–57). Wiley Blackwell.
- Madrid-Morales, D., Wasserman, H., Gondwe, G., Ndlovu, K., Sikanku, E., Tully, M., Umejei, E., & Uziegbunam, C. (2021). Motivations for sharing misinformation: A comparative study in six sub-Saharan African countries. *International Journal of Communication*, 15, 1200–1219. <https://ijoc.org/index.php/ijoc/article/view/14801>
- Magalhães, J. C. (2018). Do algorithms shape character? Considering algorithmic ethical subjectivation. *Social Media + Society*, 4(2), 1–10. <https://doi.org/10.1177/2056305118768301>
- Magalhães, J. C., & Couldry, N. (2021). Giving by taking away: Big tech, data colonialism, and the reconfiguration of social good. *International Journal of Communication*, 15, 343–362. <https://ijoc.org/index.php/ijoc/article/view/15995>
- Mager, A., & Katzenbach, C. (2021). Future imaginaries in the making and governing of digital technology: Multiple, contested, commodified. *New Media & Society*, 23(2), 223–236. <https://doi.org/10.1177/1461444820929321>
- Mahler, T. (2022). Between risk management and proportionality: The risk-based approach in the EU’s artificial intelligence act proposal. In L. Colonna, & S. Greenstein (Eds.), *Law in the era of artificial intelligence: Nordic yearbook of law and informatics 2020-2021* (pp. 247–270). The Swedish Law and Informatics Research Institute.

- Mahomed, S., Aitken, M., Atabey, A., Wong, J., & Briggs, M. (2023). *AI, children's rights, & wellbeing: Transnational frameworks: Mapping 13 frameworks at the intersections of data-intensive technologies, children's rights and wellbeing*. Alan Turing Institute UK. <https://www.turing.ac.uk/news/publications/ai-childrens-rights-wellbeing-transnational-frameworks>
- Majal. (2024a). *About us*. Retrieved July 5, 2025 from <https://majal.org/about-us>
- Majal. (2024b). *Numun Fund: Supporting feminist tech infrastructure*. Retrieved July 5, 2025 from <https://majal.org/numun-fund>
- Makan, N. (2023). Sustainable industrial machine learning and artificial intelligence: A framework for environmentally responsible innovation. *Transactions on Recent Developments in Artificial Intelligence and Machine Learning*, 15(15). <https://journals.throws.com/index.php/TRDAIML/article/view/184>
- Makhortykh, M., Urman, A., Münch, F. V., Heldt, A., Dreyer, S., & Kettemann, M. C. (2022). Not all who are bots are evil: A cross-platform analysis of automated agent governance. *New Media & Society*, 24(4), 964–981. <https://doi.org/10.1177/14614448221079035>
- Malaysia Government. (2018). *Anti-fake news bill 2018*. https://www.cljlaw.com/files/bills/pdf/2018/MY_FS_BIL_2018_06.pdf
- Mano, W., & Milton, V. C. (Eds.). (2021). *Routledge handbook of African media and communication studies*. Routledge. <https://doi.org/10.4324/9781351273206>
- Mansell, R. (2012). *Imagining the internet: Communication, innovation and governance*. Oxford University Press.
- Mansell, R. (2021). European responses to (US) digital platform dominance. In D. Y. Jin (Ed.), *The Routledge handbook of digital media and globalization* (pp. 141–149). Routledge. <https://doi.org/10.4324/9780367816742>
- Mansell, R., Durach, F., Kettemann, M., Lenoir, T., Procter, R., Tripathi, G., & Tucker, E. (2025, January 15). *Information ecosystem and troubled democracy: A global synthesis of the state of knowledge on news media, ai and data governance*. Observatory on Information and Democracy. <https://observatory.informationdemocracy.org/report/information-ecosystem-and-troubled-democracy/>
- Mansell, R., & Nordenstreng, K. (2006). Great media and communications debates – An assessment of the MacBride report after 25 years. *Information Technologies and International Development*, 3(4), 15–36. <https://itidjournal.org/index.php/itid/article/view/235.html>
- Mansell, R., & Steinmueller, W. E. (2020). *Advanced introduction to platform economics*. Edward Elgar Publishing.
- Marconi, F. (2020). *Newsmakers – Artificial intelligence and the future of journalism*. Columbia University Press.
- Marcucci, S., Alarcón, N. G., Verhulst, S. G., & Wüllhorst, E. (2023). Informing the global data future: Benchmarking data governance frameworks. *Data & Policy*, 5(e30), 1–25. <https://doi.org/10.1017/dap.2023.24>
- Marda, V., & Milan, S. (2018). *Wisdom of the crowd: Multistakeholder perspectives on the fake news debate*. Internet Policy Observatory, Annenberg School for Communication, University of Pennsylvania; DATACTIVE Ideas Lab, NL. <https://papers.ssrn.com/abstract=3184458>
- Mare, A. (2020). State-ordered internet shutdowns and digital authoritarianism in Zimbabwe. *International Journal of Communication*, 14, 4244–4263. <https://ijoc.org/index.php/ijoc/article/view/11494>
- Mare, A., Mabwezara, H. M., & Moyo, D. (2019). “Fake news” and cyber-propaganda in sub-Saharan Africa: Recentring the research agenda. *African Journalism Studies*, 40(4), 1–12. <https://doi.org/10.1080/23743670.2020.1788295>
- Marinova, Z., Spangenberg, J., Teyssou, D., Papadopoulos, S., Sarris, N., Alaphilippe, A., & Bontcheva, K. (2020). Weverify: Wider and enhanced verification for your project overview and tools. In *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. <https://ieeexplore.ieee.org/document/9106056>
- Marsden, C. T. (2016). Comparative case studies in implementing net neutrality: A critical analysis of zero rating. *SCRIPTed: A Journal of Law, Technology & Society*, 13(1), 1–39. <https://doi.org/10.2966/script.130116.1>

- Marsden, C. T., & Brown, I. (2023). App stores, antitrust and their links to net neutrality: A review of the European policy and academic debate leading to the EU Digital Markets Act. *Internet Policy Review*, 12(1), 1–27. <https://policyreview.info/articles/analysis/app-stores-antitrust-net-neutrality-eu-digital-markets-act>
- Marshall, S. (2023). *We get past “post-platform”*. Nieman Lab. <https://www.niemanlab.org/2023/12/we-get-past-post-platform/>
- Martens, M., De Wolf, R., Berendt, B., & De Marez, L. (2023). Decoding algorithms: Exploring end-users’ mental models of the inner workings of algorithmic news recommenders. *Digital Journalism*, 11(1), 203–225. <https://doi.org/10.1080/21670811.2022.2129402>
- Martin, K. (2022). Google research: Who is responsible for ethics of AI? In K. Martin (Ed.), *Ethics of data and analytics: Concepts and cases* (pp. 434–446). Auerbach Publications. <https://doi.org/10.1201/9781003278290>
- Martin, A., & Newell, B. (2024). Synthetic data, synthetic media, and surveillance. *Surveillance & Society*, 22(4), 448–452. <https://doi.org/10.24908/ss.v22i4.18334>
- Martínez-Costa, M.-P., López-Pan, F., Buslón, N., & Salaverría, R. (2023). Nobody-fools-me perception: Influence of age and education on overconfidence about spotting disinformation. *Journalism Practice*, 17(10), 2084–2102. <https://doi.org/10.1080/17512786.2022.2135128>
- Masur, P. K. (2020). How online privacy literacy supports self-data protection and self-determination in the age of information. *Media and Communication*, 8(2), 1–12. <https://doi.org/10.17645/mac.v8i2.2855>
- Masur, P. K., DiFranzo, D., & Bazarova, N. N. (2021). Behavioral contagion on social media: Effects of social norms, design interventions, and critical media literacy on self-disclosure. *PLOS ONE*, 16(7). <https://doi.org/10.1371/journal.pone.0254670>
- Matasick, C., Alfonsi, C., & Bellantoni, A. (2020). *Governance responses to disinformation: How open government principles can inform policy options*. OECD. <https://www.oecd-ilibrary.org/content/paper/d6237c85-en>
- Mateen, H., Tabakovic, H., Holder, P., & Schiffrin, A. (2023). *Paying for news: What Google and Meta owe publishers*. Initiative for Policy Dialogue, Columbia University. <http://dx.doi.org/10.2139/ssrn.4704237>
- Mattelart, T. (2023). US digital platforms in the Global South: A critical review of an emerging research field. In P. Bouquillion, C. Ithurbide, & T. Mattelart (Eds.), *Digital platforms and the Global South: Reconfiguring power relations in the cultural industries* (pp. 18–36). Routledge. <https://doi.org/10.4324/9781003391746>
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work and think*. John Murray.
- Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 1906–1919). Association for Computational Linguistics.
- Mayson, S. G. (2018). Bias in, bias out. *The Yale Law Journal*, 2393, 2218–2300. https://scholarship.law.upenn.edu/faculty_scholarship/2393
- Mazzoli, E. M. (2020). Online content governance: Towards a framework for analysis for prominence and discoverability. *Journal of Digital Media & Policy*, 11, 301–319. https://doi.org/10.1386/jdmp_00027_1
- Mazzoli, E. M. (2023). *Digital trust initiatives: Seeking to reward journalistic ethics online*. Center for International Media Assistance and National Endowment for Democracy. <https://www.cima.ned.org/publication/digital-trust-initiatives/>
- Mazzucato, M. (2023). Governing the economics of the common good: From correcting market failures to shaping collective goals. *Journal of Economic Policy Reform*, 27(1), 1–24. <https://doi.org/10.1080/17487870.2023.2280969>
- McCombs, M., & Shaw, D. L. (1972). The agenda-setting function of mass media. *The Public Opinion Quarterly*, 36(2), 176–187. <https://doi.org/10.1086/267990>
- McConvey, J. R. (2024, May 22). *Police in US cities that ban facial recognition asking others to do it for them*. Biometric Update. <https://www.biometricupdate.com/202405/police-in-us-cities-that-ban-facial-recognition-asking-others-to-do-it-for-them>
- McGonagle, T., Bednarski, M., Francese Coutinho, M., & Zimin, A. (2019). *Elections and media in digital times*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000371486>

- McGuigan, L. (2023). *Selling the American people: Advertising, optimization, and the origins of adtech*. MIT Press. <https://doi.org/10.7551/mitpress/13562.001.0001>
- McIntyre, K., & Sobel, M. (2019). How Rwandan journalists use WhatsApp to advance their profession and collaborate for the good of their country. *Digital Journalism*, 7(6), 705–724. <https://doi.org/10.1080/21670811.2019.1612261>
- McKernan, B., Rossini, P., & Stromer-Galley, J. (2023). Echo chambers, cognitive thinking styles, and mistrust? Examining the roles information sources and information processing play in conspiracist ideation. *International Journal of Communication*, 17, 1102–1125. <https://ijoc.org/index.php/ijoc/article/view/19244>
- McLuhan, M. (1964). *Understanding media: The extensions of man*. McGraw-Hill.
- McQuail, D. (2010). *Mass communication theory* (6th ed.). Sage.
- Medrado, A., & Verdegem, P. (2024). Participatory action research in critical data studies: Interrogating AI from a South–North approach. *Big Data & Society*, 11(1), 1–14. <https://doi.org/10.1177/20539517241235869>
- Medronho, R. (2024, June 14). *Far-right seeks to censor science and criminalize researchers studying disinformation in Brazil*. Tech Policy Press. <https://techpolicy.press/farright-seeks-to-censor-science-and-criminalize-researchers-studying-disinformation-in-brazil>
- Meese, J., & Hurcombe, E. (2021). Facebook, news media and platform dependency: The institutional impacts of news distribution on social platforms. *New Media & Society*, 23(8), 2367–2384. <https://doi.org/10.1177/1461444820926472>
- Mehta, S., & Erickson, K. (2022). Can online political targeting be rendered transparent? Prospects for campaign oversight using the Facebook Ad Library. *Internet Policy Review*, 11(1), 1–32. <https://doi.org/10.14763/2022.1.1648>
- Mejias, U. A., & Couldry, N. (2024). *Data grab: The new colonialism of big tech and how to fight back*. The University of Chicago Press.
- Melih, Y. (2022). New geographies of platform capitalism: The case of digital monopolization in Turkey. *Big Data & Society*, 9(2), 1–14. <https://doi.org/10.1177/20539517221124585>
- Mendoza, R. U., Elemia, C. K. S., Recto, J. M. M., & de Castro, B. A. B. (2023). When fake news infects political networks: Case study of the Tallano gold myth in the Philippines. *Media Asia*, 50(4), 501–527. <https://doi.org/10.1080/01296612.2023.2217607>
- Menn, J., & Nix, N. (2023, December 6). Big tech funds the very people who are supposed to hold it accountable. *The Washington Post*. <https://www.washingtonpost.com/technology/2023/12/06/academic-research-meta-google-university-influence>
- Menon, S. (2017). The role of civil society groups in improving access to the DC-CAN (District of Columbia – Community Access Network). *Information Systems Frontiers*, 19(2), 361–375. <https://doi.org/10.1007/s10796-015-9602-1>
- Menon, S. (2021). An institutional analysis of TMP regulation in India. *Review of Policy Research*, 38(3), 300–325. <https://doi.org/10.1111/ropr.12425>
- Menon, S. (2024). The digital navigator programme in the time of COVID-19: A case study on Philadelphia’s programme. *Journal of Digital Media & Policy*, 1–26. Online First. https://doi.org/10.1386/jdmp_00140_1
- Merenda, F. (2021). Reading Arendt to rethink truth, science and politics in the era of fake news. In S. Giusti, & E. Piras (Eds.), *Democracy and fake news: Information manipulation and post-truth politics* (pp. 19–29). Routledge. <https://doi.org/10.4324/9781003037385>
- Metz, C. (2023, March 31). The ChatGPT king isn’t worried, but he knows you might be. *The New York Times*. <https://www.nytimes.com/2023/03/31/technology/sam-altman-open-ai-chatgpt.html>
- Metzger, M. J., Hartsell, E. H., & Flanagin, A. J. (2020). Cognitive dissonance or credibility? A comparison of two theoretical explanations for selective exposure to partisan news. *Communication Research*, 47(1), 3–28. <https://doi.org/10.1177/0093650215613136>
- Michael, A. (2023). *Artificial intelligence, democracy and elections*. European Parliament Briefing. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/751478/EPRS_BRI\(2023\)751478_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/751478/EPRS_BRI(2023)751478_EN.pdf)
- Michalis, M., & D’Arma, A. (2024). Public service media: From epistemic rights to justice. In M. A. Horowitz, H. Nieminen, K. Lehtisaari, & A. D’Arma (Eds.), *Epistemic rights in the era of digital disruption* (pp. 97–109). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-45976-4_7

- Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society*, 7(2), 1–15. <https://doi.org/10.1177/2053951720948087>
- Microsoft. (2023). *Diversity and inclusion report 2023: A decade of transparency, commitment and progress*. <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW1e53b>
- Miguel, R., & Krack, N. (2023, September 28). *Platforms' policies on AI-manipulated and generated misinformation*. EU DisinfoLab. <https://www.disinfo.eu/publications/platforms-policies-on-ai-manipulated-and-generated-misinformation>
- Milan, S., & Treré, E. (2019). Big data from the South(s): Beyond data universalism. *Television & New Media*, 20(4), 319–335. <https://doi.org/10.1177/1527476419837739>
- Milan, S., & van der Velden, L. (2016). The alternative epistemologies of data activism. *Digital Culture & Society*, 2(2), 57–74. <https://doi.org/10.14361/dcs-2016-0205>
- Milan, S., Treré, E., & Masiero, S. (Eds.). (2021). *COVID-19 from the margins: Pandemic invisibilities, policies and resistance in the datafied society*. Institute of Network Cultures, Amsterdam University.
- Milmo, D. (2024, May 28). OpenAI putting 'shiny products' above safety, says departing researcher. *The Observer*. <https://www.theguardian.com/technology/article/2024/may/18/openai-putting-shiny-products-above-safety-says-departing-researcher>
- Miltner, K., & Highfield, T. (2024). *The possibilities of 'good' generative AI in the cultural and creative industries*. The British Academy. <https://www.thebritishacademy.ac.uk/publications/the-possibilities-of-good-generative-ai-in-the-cultural-and-creative-industries/>
- milton, v. c., & Mano, W. (2022). Afrokology and the right to communicate in Africa. *Javnost – The Public*, 29(1), 33–49. <https://doi.org/10.1080/13183222.2021.1889830>
- Ministry of Local Government and Regional Development. (2025). *Artificial intelligence and democratic elections – International experiences and national recommendations report by the expert group on artificial intelligence and elections*. Norwegian Government. <https://www.regjeringen.no/en/dokumenter/artificial-intelligence-and-democratic-elections-international-experiences-and-national-recommendations-report-by-the-expert-group-on-artificial-intelligence-and-elections/id3086085/>
- Misra, S. (2022, March 5). *The government wants to sell your data* | #SaveOurPrivacy. Internet Freedom Foundation. <https://internetfreedom.in/the-government-wants-to-sell-your-data>
- MIT Technology Review Insights. (2023). *Finding value in generative AI for financial services*. MIT Technology Review; UBS. <https://www.technologyreview.com/2023/11/26/1083841/finding-value-in-generative-ai-for-financial-services>
- Mitchell, S., Potash, E., Barocas, S., D'Amour, A., & Lum, K. (2021). Algorithmic fairness: Choices, assumptions, and definitions. *Annual Review of Statistics and Its Application*, 8, 141–163. <https://doi.org/10.1146/annurev-statistics-042720-125902>
- Mitchelstein, E., & Boczkowski, P. J. (Eds.). (2023). *Digital journalism in Latin America*. Routledge. <https://doi.org/10.4324/9781003370383>
- Mitchelstein, E., Matassi, M., & Boczkowski, P. J. (2020). Minimal effects, maximum panic: Social media and democracy in Latin America. *Social Media + Society*, 6(4), 1–11. <https://doi.org/10.1177/2056305120984452>
- Mittelstadt, B., Wachter, S., & Russell, C. (2023). To protect science, we must use LLMs as zero-shot translators. *Nature Human Behaviour*, 7(11), 1830–1832. <https://doi.org/10.1038/s41562-023-01744-0>
- Mitts, T., Phillips, G., & Walter, B. F. (2022). Studying the impact of ISIS propaganda campaigns. *The Journal of Politics*, 84(2), 1220–1225. <https://doi.org/10.1086/716281>
- Moens, B. (2025, May 16). EU moves to settle Microsoft Teams antitrust case. *Financial Times*. <https://www.ft.com/content/3c6548b1-97b6-4891-acae-75b776ef2cb3>
- Mökander, J., Schuett, J., Kirk, H. R., & Floridi, L. (2023). Auditing large language models: A three-layered approach. *AI and Ethics*, 4, 1085–1115. <https://doi.org/10.1007/s43681-023-00289-2>
- Molina, M. D., & Sundar, S. S. (2022). Does distrust in humans predict greater trust in AI? Role of individual differences in user responses to content moderation. *New Media & Society*, 26(6), 3638–3656. <https://doi.org/10.1177/14614448221103534>

- Möller, J., Hameleers, M., & Ferreau, F. (2020). *Types of disinformation and misinformation: Various types of disinformation and their dissemination from a communication science and legal perspective*. Committee Chairperson Conference (GVK, Gremienvorsitzendenkonferenz) of the German Media Authorities. https://www.die-medienanstalten.de/fileadmin/user_upload/die_medienanstalten/Service/Studien_und_Gutachten/GVK_Summary_EN_final_web.pdf
- Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame it on the algorithm: An empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society*, 21(7), 959–977. <https://doi.org/10.1080/1369118X.2018.1444076>
- Monasterio Astobiza, A., Ausín, T., Liedo, B., Toboso, M., Aparicio, M., & López, D. (2022). Ethical governance of AI in the Global South: A human rights approach to responsible use of AI. *Proceedings* 81(136), 1–5. <https://doi.org/10.3390/proceedings2022081136>
- Mont’Alverne, C., Badrinathan, S., Ross Arguedas, A., Toff, B., Fletcher, R., & Nielsen, R. (2022). *The trust gap: How and why news on digital platforms is viewed more sceptically versus news in general*. Reuters Institute for the Study of Journalism, University of Oxford. <https://ora.ox.ac.uk/objects/uuid:42cc0bd8-f737-4a79-947f-e528e8116926>
- Mont’Alverne, C., Ross Arguedas, A., Badrinathan, S., Toff, B., Fletcher, R., & Nielsen, R. K. (2023). Domain-specific influence on Facebook: How topic matters when assessing influential accounts in four countries. *Journal of Quantitative Description: Digital Media*, 3, 1–34. <https://dx.doi.org/10.51685/jqd.2023.014>
- Mont’Alverne, C., Ross Arguedas, A., Banerjee, S., Toff, B., Fletcher, R., & Nielsen, R. (2024). The electoral misinformation nexus: How news consumption, platform use, and trust in news influence belief in electoral misinformation. *Public Opinion Quarterly*, 88(S1), 681–707. <https://doi.org/10.1093/poq/nfae019>
- Montero, J., & Finger, M. (2021). *The rise of the new network industries: Regulating digital platforms*. Routledge. <https://doi.org/10.4324/9781003141327>
- Monti, M. (2020). The EU code of practice on disinformation and the risk of the privatisation of censorship. In S. Giusti, & E. Piras (Eds.), *Democracy and fake news: Information manipulation and post-truth politics* (pp. 214–225). Routledge. <https://doi.org/10.4324/9781003037385>
- Monzer, C., Moeller, J., Helberger, N., & Eskens, S. (2020). User perspectives on the news personalisation process: Agency, trust and utility as building blocks. *Digital Journalism*, 8(9), 1142–1162. <https://doi.org/10.1080/21670811.2020.1773291>
- Moore, M., & Tambini, D. (Eds.). (2018). *Digital dominance: The power of Google, Amazon, Facebook, and Apple*. Oxford University Press.
- Moore, M., & Tambini, D. (Eds.). (2021). *Regulating big tech: Policy responses to digital dominance*. Oxford University Press. <https://doi.org/10.1093/oso/9780197616093.001.0001>
- Moore, R. C., & Hancock, J. T. (2022). A digital media literacy intervention for older adults improves resilience to fake news. *Nature Scientific Reports*, 12(1), 1–9. <https://doi.org/10.1038/s41598-022-08437-0>
- Moran, C. R., Burton, J., & Christou. (2023). The US intelligence community, global security and AI: From secret intelligence to smart spying. *Journal of Global Security Studies*, 8(2), 1–18. <https://doi.org/10.1093/jogss/ogad005>
- Moran, R. E., & Shaikh, S. J. (2022). Robots in the news and newsrooms: Unpacking Meta-journalistic discourse on the use of artificial intelligence in journalism. *Digital Journalism*, 10(10), 1756–1774. <https://doi.org/10.1080/21670811.2022.2085129>
- Moreno Belloso, N., & Petit, N. (2023). The EU Digital Markets Act (DMA): A competition hand in a regulatory glove. *European Law Review*, 48(4), 391–421. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4411743
- Morkūnas, M. (2023). Russian disinformation in the Baltics: Does it really work? *Public Integrity*, 25(6), 599–613. <https://doi.org/10.1080/10999922.2022.2092976>
- Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism*. Public Affairs.
- Motamedi, M. (2024, February 24). *Iran unveils plan for tighter internet rules to promote local platforms*. Al Jazeera. <https://www.aljazeera.com/news/2024/2/24/iran-unveils-plan-for-tighter-internet-rules-to-promote-local-platforms>

- Motlagh, N. Y., Khajavi, M., Sharifi, A., & Ahmadi, M. (2023). The impact of artificial intelligence on the evolution of digital education: A comparative study of OpenAI text generation tools including ChatGPT, Bing Chat, Bard, and Ernie. *ArXiv*. <https://arxiv.org/abs/2309.02029>
- Motta, M., Hwang, J., & Stecula, D. (2024). What goes down must come up? Pandemic-related misinformation search behavior during an unplanned Facebook outage. *Health Communication*, 39(10), 2041–2052. <https://doi.org/10.1080/10410236.2023.2254583>
- Mudde, C. (2019). *The far right today*. Polity Press.
- Mouffe, C. (1966). Democracy, power, and the ‘political’. In S. Benhabib (Ed.), *Democracy and difference: Contesting the boundaries of the political* (pp. 245–256). Princeton University Press.
- Mueller, M. L. (2025). It’s just distributed computing: Rethinking AI governance. *Telecommunications Policy*, 49(3), 1–12. <https://doi.org/10.1016/j.telpol.2025.102917>
- Mueller, M. (2024, August 27). *The Justice Department sues Georgia Tech: A teachable moment in cybersecurity management*. Internet Governance Project. <https://www.internetgovernance.org/2024/08/27/the-justice-department-sues-georgia-tech-a-teachable-moment-in-cybersecurity-management>
- Mühlhoff, R., & Ruschmeier, H. (2024). Regulating AI with purpose limitation for models. *Journal of AI Law and Regulation*, 1(1), 24–39. <https://doi.org/10.21552/aire/2024/1/5>
- Mukerjee, S. (2016). Net neutrality, Facebook, and India’s battle to #SaveTheInternet. *Communication and the Public*, 1(3), 356–361. <https://doi.org/10.1177/2057047316665850>
- Mukhudwana, R. F. (2021). The rise of peripheral actors in media regulation in South Africa: An entry of social media mob(s). *African Journalism Studies*, 42(4), 153–178. <https://doi.org/10.1080/23743670.2022.2032783>
- Müller, M., & Kettemann, M. C. (2024). European approaches to the regulation of digital technologies. In H. Werthner, C. Ghezzi, J. Kramer, J. Nida-Rümelin, B. Nuseibeh, E. Prem, & A. Stanger (Eds.), *Introduction to digital humanism: A textbook* (pp. 623–637). Springer. <https://doi.org/10.1007/978-3-031-45304-5>
- Mulligan, C. E. A., & Godsiff, P. (2023). Datalism and data monopolies in the era of AI: A research agenda. *arXiv:2307.08049v1*. <https://doi.org/10.48550/arXiv.2307.08049>
- Munoriyarwa, A., & Mare, A. (2022). Mainstreaming surveillance through the biometrification of everyday life. In A. Munoriyarwa, & A. Mare (Eds.), *Digital surveillance in southern Africa: Policies, politics and practices* (pp. 141–156). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-16636-5_5
- Muñoz, K. (2023). *The transformative role of AI in reshaping electoral politics* [DGAP Memo, No. 4]. Deutsche Gesellschaft für Auswärtige Politik e.V. https://dgap.org/system/files/article_pdfs/DGAP-MEMO-No.04-September-2023_3pp_0.pdf
- Murgia, M. (2019). Google accused of secretly feeding personal data to advertisers. *The Financial Times*. <https://www.ft.com/content/e3e1697e-ce57-11e9-99a4-b5ded7a7fe3f>
- Murphy, H., Pooler, M., & Harris, B. (2024, September 19). Brazil imposes fine on Musk’s X after service returns. *The Financial Times*. <https://www.ft.com/content/2133cca8-ff48-4e11-8ca7-bd35d69201c5>
- Murtaza, G., & Salman, S. (2019). *GDPR fine imposed upon Google: An analysis*. Manuscript. Brown University. <https://cs.brown.edu/courses/csci2390/2019/assign/gdpr/ssalman1-gmurtaza-google.pdf>
- Musiani, F. (2022). Infrastructuring digital sovereignty: A research agenda for an infrastructure-based sociology of digital self-determination practices. *Information, Communication & Society*, 25(6), 785–800. <https://doi.org/10.1080/1369118X.2022.2049850>
- Musoni, M., Karkare, P., & Teevan, C. (2024). *Cross-border data flows in Africa: Continental ambitions and political realities* [Discussion Paper No. 379]. European Centre for Development Policy Management. <https://ecdpm.org/work/cross-border-data-flows-africa-continental-ambitions-and-political-realities>
- Mutsaers, B. (2019). Challenges facing development of data journalism in non-Western societies. *Digital Journalism*, 7(9), 1289–1294. <https://doi.org/10.1080/21670811.2019.1691927>

- Mutsvairo, B., & Bebawi, S. (2019). Journalism educators, regulatory realities, and pedagogical predicaments of the “fake news” era: A comparative perspective on the Middle East and Africa. *Journalism & Mass Communication Educator*, 74(2), 143–157. <https://doi.org/10.1177/1077695819833552>
- Mutsvairo, B., & Ragnedda, M. (Eds.). (2019). *Mapping digital divide in Africa: A mediated analysis*. Amsterdam University Press.
- Mutsvairo, B., & Rønning, H. (2020). The Janus face of social media and democracy? Reflections on Africa. *Media, Culture & Society*, 42(3), 317–328. <https://doi.org/10.1177/0163443719899615>
- Myers, M., & Gilberds, H. (2024). *Are donors taking the journalism crisis seriously? An analysis of official aid to media 2010–2019*. Center for International Media Assistance, National Endowment for Democracy. <https://www.cima.ned.org/publication/are-donors-taking-the-journalism-crisis-seriously/>
- Naem, S. B., Bhatti, R., & Khan, A. (2021). An exploration of how fake news is taking over social media and putting public health at risk. *Health Information & Libraries Journal*, 38(2), 143–149. <https://doi.org/10.1111/hir.12320>
- Nagaraj, A. (2020, March 17). *Under watch: Indian city workers protest digital surveillance*. Reuters. <https://www.reuters.com/article/idUSKBN21404S>
- Nahmias, Y., & Perel, M. (2021). The oversight of content moderation by AI: Impact assessments and their limitations. *Harvard Journal on Legislation*, 58(1), 145–194. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3565025
- Nannini, L., Bonel, E., Bassi, D., & Maggini, M. J. (2024). Beyond phase-in: Assessing impacts on disinformation of the EU digital services act. *AI and Ethics*, 5, 1241–1269. <https://doi.org/10.1007/s43681-024-00467-w>
- Napoli, P. M. (2003). *Audience economics: Media institutions and the audience marketplace*. Columbia University Press.
- Napolitano, A. (2023). *Artificial intelligence: The new frontier of the EU’s border externalisation strategy*. EuroMed Rights. https://euromedrights.org/wp-content/uploads/2023/07/Euromed_AI-Migration-Report_EN-1.pdf
- Nardi, B. A., & O’Day, V. (1999). *Information ecologies: Using technology with heart*. MIT Press.
- Ncube, L. (2019). Digital media, fake news and pro-movement for democratic change (MDC) alliance cyber-propaganda during the 2018 Zimbabwe election. *African Journalism Studies*, 40(4), 44–61. <https://doi.org/10.1080/23743670.2019.1670225>
- Ndemo, B., & Thegeya, A. (2023). A prototype data governance framework for Africa. In B. Ndemo, N. Ndung’u, S. Odhiambo, & A. Shimeles (Eds.), *Data governance and policy in Africa* (pp. 9–29). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-24498-8_2
- NDI (National Democratic Institute). (2023). *Mitigating online misleading information and polarization in conflict-sensitive contexts: Experimental evidence from Côte d’Ivoire*. <https://www.ndi.org/publications/mitigating-online-misleading-information-and-polarization-conflict-sensitive-contexts>
- NDIA (National Digital Inclusion Alliance). (n.d.). *About NDIA*. <https://www.digitalinclusion.org/about-ndia>
- Nechushtai, E. (2018). Could digital platforms capture the media through infrastructure? *Journalism*, 19(8), 1043–1058. <https://doi.org/10.1177/1464884917725163>
- Nechushtai, E. (2023). Resisting the individualization of risk: Strategies of engagement and caution in journalists’ responses to online mobs in the United States and Germany. *Digital Journalism*, 11(10), 1906–1923. <https://doi.org/10.1080/21670811.2022.2153710>
- Neff, T., & Pickard, V. (2024). Funding democracy: Public media and democratic health in 33 countries. *The International Journal of Press/Politics*, 29(3), 601–627. <https://doi.org/10.1177/19401612211060255>
- Neenadić, I., Brogi, E., & Bleyer-Simon, K. (2023). *Structural indicators to assess effectiveness of the EU’s code of practice on disinformation* [RSC Working Paper 2023/34]. European University Institute. <https://cadmus.eui.eu/handle/1814/75558>
- Nettle, D. (2023, June 14). *The political economy of scientific publishing, and the promise of diamond open access* [Blog post]. <https://www.danielnettle.org.uk/2023/06/14/the-political-economy-of-scientific-publishing-and-the-promise-of-diamond-open-access>

- Neubaum, G., & Weeks, B. (2023). Computer-mediated political expression: A conceptual framework of technological affordances and individual tradeoffs. *Journal of Information Technology & Politics*, 20(1), 19–33. <https://doi.org/10.1080/19331681.2022.2028694>
- Newman, N., & Fletcher, R. (2017). *Bias, bullshit and lies: Audience perspectives on low trust in the media: Digital news project 2017*. Reuters Institute for the Study of Journalism, University of Oxford. <https://reutersinstitute.politics.ox.ac.uk/our-research/bias-bullshit-and-lies-audience-perspectives-low-trust-media>
- Newman, N., Fletcher, R., Robertson, C. T., Kirsten, E., & Nielsen, R. K. (2022). *Digital news report 2022*. Reuters Institute for the Study of Journalism, University of Oxford. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022>
- Newman, N., Fletcher, R., Eddy, K., Robinson, C. T., & Nielsen, R. K. (2023). *Digital news report 2023*. Reuters Institute for the Study of Journalism, University of Oxford. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023>
- Newman, N., Fletcher, R., Craig, T. R., Ross Arguedas, A., & Nielsen, R. K. (2024). *Digital news report 2024*. Reuters Institute for the Study of Journalism, University of Oxford. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024>
- Newton, C., & Schiffer, Z. (2024, June 13). *The Stanford internet observatory is being dismantled*. Platformer. <https://www.platformer.news/stanford-internet-observatory-shutdown-stamos-diresta-sio>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). AI literacy: Definition, teaching, evaluation and ethical issues. *Proceedings of the Association for Information Science and Technology*, 58(1), 504–509. <https://doi.org/10.1002/pra2.487>
- Nguyễn, S., Kuo, R., Reddi, M., Li, L., & Moran, R. E. (2022). Studying mis- and disinformation in Asian diasporic communities: The need for critical transnational research beyond Anglocentrism. *Harvard Kennedy School (HKS) Misinformation Review*, 3(2), 1–12. <https://doi.org/10.37016/mr-2020-95>
- Nicas, J., & Conger, K. (2024, August 30). Brazil blocks X after Musk ignores court orders. *The New York Times*. <https://www.nytimes.com/2024/08/30/world/americas/brazil-elon-musk-x-blocked.html>
- Nicholas, G., & Bhatia, A. (2023). *Lost in translation: Large language models in non-English content analysis*. Center for Democracy & Technology. <https://cdt.org/insights/lost-in-translation-large-language-models-in-non-english-content-analysis/>
- Nicholson, C. (2024, February 15). *Elsevier parent reports 10% hike in profits for 2023*. Research Professional News. <https://www.researchprofessionalnews.com/tr-news-europe-infrastructure-2024-2-elsevier-parent-reports-10-hike-in-profits-for-2023>
- Nicoli, N., & Iosifidis, P. (2023). EU digital economy competition policy: From ex-post to ex-ante. The case of Alphabet, Amazon, Apple, and Meta. *Global Media and China*, 8(1), 24–38. <https://doi.org/10.1177/20594364231152673>
- Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 20(11), 4275–4292. <https://doi.org/10.1177/1461444818769694>
- Nieborg, D., Poell, T., Caplan, R., & van Dijck, J. (2024). Locating and theorising platform power. *Internet Policy Review*, 13(2), 1–17. <https://policyreview.info/articles/analysis/introduction-special-issue-locating-and-theorising-platform-power>
- Nielsen, R. K., & Fletcher, R. (2023). Comparing the platformization of news media systems: A cross-country analysis. *European Journal of Communication*, 38(5), 484–499. <https://doi.org/10.1177/02673231231189043>
- Nielsen, R. K., & Ganter, S. A. (2022). *The power of platforms: Shaping media and society*. Oxford University Press. <https://doi.org/10.1093/oso/9780190908850.001.0001>
- Nielsen, R. K., Gorwa, R., & de Cock Buning, M. (2019). *What can be done? Digital media policy options for strengthening European democracy*. Reuters Institute for the Study of Journalism, University of Oxford. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-11/What_Can_Be_Done_FINAL.pdf
- Nieminen, H. (2024). Why we need epistemic rights. In M. A. Horowitz, H. Nieminen, K. Lehtisaari, & A. D'Arma (Eds.), *Epistemic rights in the era of digital disruption* (pp. 11–28). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-45976-4_2

- Niklas, J., & Dencik, L. (2024). Data justice in the “twin objective” of market and risk: How discrimination is formulated in EU’s AI policy. *Policy & Internet*, 16(3), 509–522.
<https://doi.org/10.1002/poi3.392>
- Ning, L., Liu, L., Wu, J., Wu, N., Berlowitz, D., Prakash, S., Green, B., O’Banion, S., & Xie, J. (2024). User-LLM: Efficient LLM contextualization with user embeddings. *arXiv*.
<https://doi.org/10.48550/arXiv.2402.13598>
- Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, 140(4), 32–48.
https://doi.org/10.1162/DAED_a_00113
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- Nordic Council of Ministers Secretariat. (2023). *A Nordic approach to democratic debate in the age of big tech : Recommendations from the Nordic think tank for tech and democracy*.
<http://dx.doi.org/10.6027/nord2023-004>
- Nothias, T. (2020). Access granted: Facebook’s free basics in Africa. *Media, Culture & Society*, 42(3), 329–348. <https://doi.org/10.1177/0163443719890530>
- Notley, T., Dezuanni, M., Zhong, H., & Chambers, S. (2020). *News and young Australians in 2020: How young people access, perceive and are affected by news media*. Western Sydney University and QUT Digital Media Research Centre.
<https://researchdirect.westernsydney.edu.au/islandora/object/uws%3A56697/>
- Novet, J. (2023, April 8). *Microsoft’s \$13 billion bet on OpenAI carries huge potential along with plenty of uncertainty*. CNBC. <https://www.cnbc.com/2023/04/08/microsofts-complex-bet-on-openai-brings-potential-and-uncertainty.html>
- Nowotny, H. (2021). *In AI we trust: Power, illusion and control of predictive algorithms*. Polity Press.
- NSF (National Science Foundation). (2023). *NSF public access plan 2.0: Ensuring open, immediate and equitable access to national science foundation funded research*.
<https://www.nsf.gov/pubs/2023/nsf23104/nsf23104.pdf>
- Nyalety, E., Parizi, R. M., Zhang, Q., & Choo, K. K. R. (2019, July 1). BlockIPFS – Blockchain-enabled interplanetary file system for forensic and trusted data traceability. *2019 IEEE International Conference on Blockchain*. <https://ieeexplore.ieee.org/document/8946164>
- Nyamnjoh, F. B. (2006). *Insiders and outsiders: Citizenship and xenophobia in contemporary southern Africa*. Zed Books.
- Nyhan, B., Settle, J., Thorson, E., Wojcieszak, M., Barberá, P., Chen, A. Y., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Guess, A. M., Kennedy, E., Kim, Y. M., Lazer, D., Malhotra, N., Moehler, D., ... Tucker, J. A. (2023). Like-minded sources on Facebook are prevalent but not polarizing. *Nature*, 620, 137–144. <https://doi.org/10.1038/s41586-023-06297-w>
- NYPD (New York City Police Department). (2024). *Public oversight of surveillance technology (POST) act impact and use policies*.
<https://www.nyc.gov/site/nypd/about/about-nypd/policy/post-act.page>
- Ó Fathaigh, R., Helberger, N., & Appelman, N. (2021). The perils of legally defining disinformation. *Internet Policy Review*, 10(4), 1–25.
<https://policyreview.info/articles/analysis/perils-legally-defining-disinformation>
- Ó Siochrú, S., Singh, P. J., & Burch, S. (2024). Just net coalition: Seeking anchors of equity and social justice in a digital world. In C. Padovani, V. Wavre, A. Hintz, G. Goggin, & P. Iosifidis (Eds.), *Global communication governance at the crossroads* (pp. 239–257). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-29616-1_14
- O’Brien, I. (2024, September 15). Data center emissions probably 662% higher than big tech claims. Can it keep up the ruse? *The Guardian*.
<https://www.theguardian.com/technology/2024/sep/15/data-center-gas-emissions-tech>
- O’Brien, M., & Swenson, A. (2024, February 16). Tech companies sign accord to combat AI-generated election trickery. *The Associated Press*.
<https://apnews.com/article/ai-generated-election-deepfakes-munich-accord-meta-google-microsoft-tiktok-x-c40924ffc68c94fac74fa994c520fc06>
- O’Connor, C. (2022, March 2). #Propaganda: Russia state-controlled media flood TikTok with Ukraine disinformation. *ISD Digital Dispatches*.
https://www.isdglobal.org/digital_dispatches/propaganda-russia-state-controlled-media-flood-tiktok-with-ukraine-disinformation

- O'Faolain, A. (2024, June 24). Court dismisses claim data watchdog failed to investigate alleged breach by Google. *The Irish Times*. <https://www.irishtimes.com/crime-law/courts/2024/06/24/court-dismisses-claim-data-watchdog-failed-to-investigate-alleged-breach-by-google>
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Harvard University Press.
- O'Neill, O. (2002). *A question of trust: The BBC Reith lectures 2002*. Cambridge University Press.
- Obreja, D. M. (2023). Mapping the political landscape on social media using bibliometrics: A longitudinal co-word analysis on Twitter and Facebook publications published between 2012 and 2021. *Social Science Computer Review*, 41(5), 1712–1728. <https://doi.org/10.1177/08944393221117749>
- OECD (Organisation for Economic Co-operation and Development). (2022a). *Going digital to advance data governance for growth and well-being*. <https://doi.org/10.1787/e3d783b0-en>
- OECD (Organisation for Economic Co-operation and Development). (2022b). *Misinformation and disinformation: An international effort using behavioural science to tackle the spread of misinformation* [OECD Public Governance Policy Papers No. 21]. <https://doi.org/10.1787/b7709d4f-en>
- OECD (Organisation for Economic Co-operation and Development). (2022c). *Recommendation of the council on artificial intelligence*. <https://oecd.ai/en/assets/files/OECD-LEGAL-0449-en.pdf>
- OECD (Organisation for Economic Co-operation and Development). (2023). *Catalogue of tools & metrics for trustworthy AI*. <https://oecd.ai/en/catalogue>
- OECD (Organisation for Economic Co-operation and Development). (2024a). *Facts not fakes: Tackling disinformation, strengthening information integrity*. https://www.oecd.org/en/publications/facts-not-fakes-tackling-disinformation-strengthening-information-integrity_d909ff7a-en.html
- OECD (Organisation for Economic Co-operation and Development). (2024b). *Recommendation of the council on information integrity*. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0505>
- Oeldorf-Hirsch, A., & Neubaum, G. (2025). What do we know about algorithmic literacy? The status quo and a research agenda for a growing field. *New Media & Society*, 27(2), 681–701. <https://doi.org/10.1177/14614448231182662>
- Ofcom. (2022). *Adults' media use and attitudes report 2022*. <https://www.ofcom.org.uk/media-use-and-attitudes/media-habits-children/children-and-parents-media-use-and-attitudes-report-2022>
- Ofcom. (2023a). *Adults' media use and attitudes report 2023*. <https://tinyurl.com/2pbsjs2u>
- Ofcom. (2023b). *Media literacy by design: Best practice principles for on-platform interventions to promote media literacy*. <https://tinyurl.com/mtb92eb8>
- Ofcom. (2023c). *News consumption in the UK: 2023 research findings*. https://www.ofcom.org.uk/_data/assets/pdf_file/0024/264651/news-consumption-2023.pdf
- Ofcom. (2023d). *Online nation 2023 report*. <https://www.ofcom.org.uk/media-use-and-attitudes/online-habits/online-nation>
- Ofcom. (2024a). *Children's media lives 2024: Ten Years of longitudinal research, a report by Ofcom*. <https://www.ofcom.org.uk/media-use-and-attitudes/media-habits-children/childrens-media-lives>
- Ofcom. (2024b). *Future technology and media literacy: Understanding generative AI*. <https://www.ofcom.org.uk/media-use-and-attitudes/media-literacy/discussion-papers/?language=en>
- Offenhuber, D. (2024). Shapes and frictions of synthetic data. *Big Data & Society*, 11(2), 1–16. <https://doi.org/10.1177/20539517241249390>
- OHCHR (Office of the United Nations High Commissioner for Human Rights). (2021). *Disinformation and freedom of opinion and expression* [Report of the special rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan, A/HRC/47/25]. <https://www.ohchr.org/en/documents/thematic-reports/ahrc4725-disinformation-and-freedom-opinion-and-expression-report>
- OHCHR (Office of the United Nations High Commissioner for Human Rights). (1993). *Vienna declaration and programme of action*. <https://www.ohchr.org/en/instruments-mechanisms/instruments/vienna-declaration-and-programme-action>

- OHCHR (Office of the United Nations High Commissioner for Human Rights). (2014). *Factors that impede equal political participation and steps to overcome those challenges*. <https://www.ohchr.org/en/documents/thematic-reports/ahrc2729-factors-impede-equal-political-participation-and-steps-overcome>
- Ohme, J., Araujo, T., Boeschoten, L., Freelon, D., Ram, N., Reeves, B. B., & Robinson, T. N. (2024). Digital trace data collection for social media effects research: APIs, data donation, and (screen) tracking. *Communication Methods and Measures*, 18(2), 124–141. <https://doi.org/10.1080/19312458.2023.2181319>
- Okolo, C. T. (2023, November 1). AI in the Global South: Opportunities and challenges towards more inclusive governance [Brookings Institution Commentary]. <https://www.brookings.edu/articles/ai-in-the-global-south-opportunities-and-challenges-towards-more-inclusive-governance>
- Okunlaya, R. O., Syed Abdullah, N., & Alias, R. A. (2022). Artificial intelligence (AI) library services innovative conceptual framework for the digital transformation of university education. *Library Hi Tech*, 40(6), 1869–1892. <https://doi.org/10.1108/LHT-07-2021-0242>
- Okyere-Manu, B. (2023). The place of the African relational and moral theory of Ubuntu in the global artificial intelligence and big data discussion: Critical reflections. In M. Findlay, L. M. Ong, & W. Zhang (Eds.), *Elgar companion to regulating AI and big data in emerging economies* (pp. 99–114). Edward Elgar Publishing. <https://doi.org/10.4337/9781785362408>
- Oladapo, O., & Ojebode, A. (2021). *Nigeria digital rights landscape report*. Institute for Development Studies. https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/15964/Nigeria_Report.pdf
- Oliver, P., Lim, C., Matthews, M., & Hanna, A. (2022). Black protests in the United States, 1994 to 2010. *Sociological Science*, 9, 275–312. <https://doi.org/10.15195/v9.a12>
- Olorunju, N., & Adams, R. (2024). African data trusts: New tools towards collective data governance? *Information & Communications Technology Law*, 33(1), 85–98. <https://doi.org/10.1080/13600834.2023.2260678>
- Omar, N. (2023, November 22). *Data, AI and elections in Africa: Working towards a collaborative African response*. Research ICT Africa. <https://researchictafrica.net/2023/11/22/data-ai-and-elections-in-africa-working-towards-a-collaborative-african-response>
- Ong, J. C. L., Chang, S. Y.-H., William, W., Butte, A. J., Shah, N. H., Chew, L. S. T., Liu, N., Doshi-Velez, F., Lu, W., Savulescu, J., & Ting, D. S. W. (2024). Ethical and regulatory challenges of large language models in medicine. *The Lancet Digital Health*, 6(6), e428–e432. [https://doi.org/10.1016/S2589-7500\(24\)00061-X](https://doi.org/10.1016/S2589-7500(24)00061-X)
- Ong, J. C., & Das, R. (2020). Two concepts from television audience research in times of datafication and disinformation. In S. Shimpach (Ed.), *The Routledge companion to global television* (pp. 111–120). Routledge. <https://doi.org/10.4324/9781315192468>
- Ong, J. C., & Tapsell, R. (2022). Demystifying disinformation shadow economies: Fake news work models in Indonesia and the Philippines. *Asian Journal of Communication*, 32(3), 251–267. <https://doi.org/10.1080/01292986.2021.1971270>
- Ortutay, B. (2025, April 29). *President Trump signs take it down act, addressing nonconsensual deepfakes. What is it?* AP News. <https://apnews.com/article/take-it-down-deepfake-trump-melania-first-amendment-741a6e525e81e5e3d8843aac20de8615>
- OSCE (Organization for Security and Co-operation in Europe). (2017, March 3). *Joint declaration on freedom of expression and “fake news”, disinformation and propaganda*. <https://www.osce.org/fom/302796>
- Our Data Indigenous. (n.d.) *Home page*. <https://ourdataindigenous.ca/>
- Ozanne, M., Bhandari, A., Bazarova, N. N., & DiFranzo, D. (2022). Shall AI moderators be made visible? Perception of accountability and trust in moderation systems on social media platforms. *Big Data & Society*, 9(2), 1–13. <https://doi.org/10.1177/20539517221115666>
- Paal, B. (2017). Current issues and recent developments on media concentration in the context of competition law and media law. *Journal of Intellectual Property Law and Practice*, 12(7), 610–616. <https://doi.org/10.1093/jiplp/jpx113>
- Padovani, C., Wavre, V., Hintz, A., Goggin, G., & Iosifidis, P. (Eds.). (2024). *Global communication governance at the crossroads*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-29616-1>

- Page, P., Knowles, P., & Mitwicki, R. (2023). Distributed governance: A principal-agent approach to data governance – Part 1. Background & core definitions. *ArXiv*.
<http://arxiv.org/abs/2308.07280>
- Palladino, N., Redeker, D., & Celeste, E. (2025). Civil society's role in constitutionalising global content governance. *Internet Policy Review*, 14(1), 1–31.
<https://doi.org/10.14763/2025.1.1830>
- Palmer, R., & Toff, B. (2022). Neither absent nor ambient: Incidental news exposure from the perspective of news avoiders in the UK, United States, and Spain. *The International Journal of Press/Politics*, 29(3), 755–773. <https://doi.org/10.1177/19401612221103144>
- Palmer, R., Toff, B., & Nielsen, R. K. (2023). Examining assumptions around how news avoidance gets defined: The importance of overall news consumption, intention, and structural inequalities. *Journalism Studies*, 24(6), 697–714.
<https://doi.org/10.1080/1461670X.2023.2183058>
- Pamment, J. (2020). *EU code of practice on disinformation: Briefing note for the new European Commission* [Working paper]. Carnegie Endowment for International Peace.
<https://policycommons.net/artifacts/430367/eu-code-of-practice-on-disinformation/1401417/>
- Pan, C. A., Yakhmi, S., Iyer, T. P., Strasnick, E., Zhang, A. X., & Bernstein, M. S. (2022). Comparing the perceived legitimacy of content moderation processes: Contractors, algorithms, expert panels, and digital juries. *Proceedings of the ACM on Human-Computer Interaction*, 6(82), 1–31. <https://doi.org/10.1145/3512929>
- Papacharissi, Z. A. (2013). *A private sphere: Democracy in a digital age*. John Wiley & Sons.
- Papaeangelou, C. (2023). Funding intermediaries: Google and Facebook's strategy to capture journalism. *Digital Journalism*, 12(2), 234–255.
<https://doi.org/10.1080/21670811.2022.2155206>
- Parcu, P. L., Brogi, E., Verza, S., Da Costa Leite Borges, D., Carlini, R., Trevisan, M., Tambini, D., Mazzoli, E. M., Klimkiewicz, B., Broughton Micova, S., Petković, B., Rossi, M. A., Stasi, M. L., Valcke, P., Lambrecht, I., Irion, K., Fahy, R., Idiz, D., ... Domazetovikj, N. (2022). *Study on media plurality and diversity online: Final report*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2759/529019>
- Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin.
- Park, G. (2020). The changing wind of data privacy law: A comparative study of the European Union's general data protection regulation and the 2018 California consumer privacy act. *UC Irvine Law Review*, 10(4), 1455–1489.
<https://escholarship.org/content/qt8562f0v0/qt8562f0v0.pdf>
- Park, S., & Humphry, J. (2019). Exclusion by design: Intersections of social, digital and data exclusion. *Information, Communication & Society*, 22(9), 934–944.
<https://doi.org/10.1080/1369118X.2019.1606266>
- Park, S., Strover, S., Choi, J., & Schnell, M. (2023). Mind games: A temporal sentiment analysis of the political messages of the internet research agency on Facebook and Twitter. *New Media & Society*, 25(3), 463–484. <https://doi.org/10.1177/14614448211014355>
- Park, Y. J. (2024). Algorithmic bias, marketplaces, and diversity regulation. *Proceedings of the TPRC 2024 Conference*, Washington DC. <https://ssrn.com/abstract=4912448>
- Parcu, P., Brogi, E., Verza, S., Da Costa Leite Borges, D., Carlini, R., Trevisan, M., Tambini, D., Mazzoli, E. M., Klimkiewicz, B., Broughton, M. S., Petković, B., Rossi, M. A., Stasi, M. L., Valcke, P., Lambrecht, I., Irion, K., Fahy, R., Idiz, D., Meiring, A., ... Domazetovikj, N. (2022) *Study on media plurality and diversity online – final report*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2759/529019>
- Pasek, J., Feldman, L., Romer, D., & Jamieson, K. H. (2008). Schools as incubators of democratic participation: Building long-term political efficacy with civic education. *Applied Developmental Science*, 12(1), 26–37. <https://doi.org/10.1080/10888690801910526>
- Pasquinelli, M. (2023). *The eye of the master: A social history of artificial intelligence*. Verso Books.
- Patrick, S., & Huggins, A. (2023, August 15). *The term “Global South” is surging. It should be retired*. Carnegie Endowment for International Peace. <https://carnegieendowment.org/posts/2023/08/the-term-global-south-is-surging-it-should-be-retired?lang=en>
- Paul, R., Carmel, E., & Cobbe, J. (Eds.). (2024). *Handbook on public policy and artificial intelligence*. Edward Elgar Publishing. <https://doi.org/10.4337/9781803922171>

- Payne, R., Martin, B. A. S., Tuzovic, S., & Wang, S. (2023). Defining biometrics with privacy and benefits: A research agenda. *Australasian Marketing Journal*, 31(4), 294–302. <https://doi.org/10.1177/14413582231167645>
- Pennycook, G., & Rand, D. G. (2022). Nudging social media toward accuracy. *The Annals of the American Academy of Political and Social Science*, 700(1), 152–164. <https://doi.org/10.1177/00027162221092342>
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7), 770–780. <https://doi.org/10.1177/0956797620939054>
- Pentney, K. (2024). States' positive obligation to create a favourable environment for participation in public debate: A principle in search of a practical effect? *Journal of Media Law*, 16(1), 146–177. <https://doi.org/10.1080/17577632.2024.2326278>
- Pequeño IV, A. (2023, July 13). Twitter starts paying creators tens of thousands – Amid intensifying competition with Threads. *Forbes*. <https://www.forbes.com/sites/antoniopequenoi/2023/07/13/twitter-starts-paying-creators-tens-of-thousands---amid-intensifying-competition-with-threads>
- Pérez-Curiel, C., Rivas-de-Roca, R., & García-Gordillo, M. (2021). Impact of Trump's digital rhetoric on the US elections: A view from worldwide far-right populism. *Social Sciences*, 10(5), 1–17. <https://doi.org/10.3390/socsci10050152>
- Perrigo, B. (2023, June 20). Exclusive: OpenAI lobbied EU to water down AI regulation. *TIME*. <https://time.com/6288245/openai-eu-lobbying-ai-act>
- Peters, J. (2016, April 5). Why is it so expensive to read academic research? *Slate*. <https://slate.com/technology/2016/04/the-lawsuit-against-sci-hub-begs-the-question-why-are-academic-journals-so-expensive-anyway.html>
- Peters, J. (2023, October 4). X stops showing headlines because Elon Musk thinks it will make posts look better. *The Verge*. <https://www.theverge.com/2023/10/4/23903859/x-elon-musk-headlines-links-image-twitter>
- Petre, C. (2021). *All the news that's fit to click: How metrics are transforming the work of journalists*. Princeton University Press.
- Pfänder, J., & Altay, S. (2025). Spotting false news and doubting true news: A systematic review and meta-analysis of news judgements. *Nature Human Behaviour*, 9, 688–699. <https://doi.org/10.1038/s41562-024-02086-1>
- Pfeiffer, J., Gutschow, J., Haas, C., Möslin, F., Maspfuhl, O., Borgers, F., & Alpsancar, S. (2023). Algorithmic fairness in AI. *Business & Information Systems Engineering*, 65(2), 209–222. <https://doi.org/10.1007/s12599-023-00787-x>
- Phan, T., Goldenfein, J., Mann, M., & Kuch, D. (2022). Economies of virtue: The circulation of 'ethics' in big tech. *Science as Culture*, 31(1), 121–135. <https://doi.org/10.1080/09505431.2021.1990875>
- Pickard, V. (2014). *America's battle for media democracy: The triumph of corporate libertarianism and the future of media reform*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139814799>
- Pickard, V. (2020a). *Democracy without journalism: Confronting the misinformation society*. Oxford University Press. <https://doi.org/10.1093/oso/9780190946753.001.0001>
- Pickard, V. (2020b). Monopoly control over digital infrastructures. In V. Pickard, *Democracy without journalism? Confronting the misinformation society* (pp. 104–135). Oxford University Press. <https://doi.org/10.1093/oso/9780190946753.001.0001>
- Pickard, V. (2020c). Restructuring democratic infrastructures: A policy approach to the journalism crisis. *Digital Journalism*, 8(6), 704–719. <https://doi.org/10.1080/21670811.2020.1733433>
- Pickard, V. (2022a). Can journalism survive in the age of platform monopolies? Confronting Facebook's negative externalities. In T. Flew, & F. R. Martin (Eds.), *Digital platform regulation: Global perspectives on internet governance* (pp. 23–42). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-95220-4_2
- Pickard, V. (2022b). Democratizing the platforms: Promises and perils of public utility regulation. *Media Development*, 68(3), 6–10. <https://tinyurl.com/2dh6bjpe>

- Pickard, V. (2024, July 16). *Legal foundations for non-reformist media reforms*. Knight First Amendment Institute, Columbia University.
<https://knightcolumbia.org/content/legal-foundations-for-non-reformist-media-reforms>
- Pickard, V., & Berman, D. E. (2019). *After net neutrality: A new deal for the digital age*. Yale University Press.
- Pielemeyer, J. (2020). Disentangling disinformation: What makes regulating disinformation so difficult? Symposium: News, disinformation, and social media responsibility. *Utah Law Review*, 2020(4), 917–940. <https://doi.org/10.26054/0D-CJBV-FTGJ>
- Pinkleton, B. E., Austin, E. W., Zhou, Y., Willoughby, J. F., & Reiser, M. (2012). Perceptions of news media, external efficacy, and public affairs apathy in political decision making and disaffection. *Journalism & Mass Communication Quarterly*, 89(1), 23–39.
<https://doi.org/10.1177/1077699011428586>
- Pintak, L., & Ginges, J. (2008). The mission of Arab journalism: Creating change in a time of turmoil. *The International Journal of Press/Politics*, 13(3), 193–227.
<https://doi.org/10.1177/1940161208317142>
- Pintak, L., & Nazir, S. J. (2013). Pakistani journalism: At the crossroads of Muslim identity, national priorities and journalistic culture. *Media, Culture & Society*, 35(5), 640–665.
<https://doi.org/10.1177/0163443713483654>
- Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293–310.
<https://doi.org/10.1177/1461444816661553>
- Pocino, P. V. (2021). *Algorithms in the newsrooms: Challenges and recommendations for artificial intelligence with the ethical values of journalism*. Catalan Press Council.
- Podolak, J., Łukasik, S., Balawender, P., Ossowski, J., Piotrowski, J., Bąkiewicz, K., & Sankowski, P. (2024). LLM generated responses to mitigate the impact of hate speech (arXiv:2311.16905). *arXiv*. <https://doi.org/10.48550/arXiv.2311.16905>
- Poell, T., Nieborg, D. B., & Duffy, B. E. (2023). Spaces of negotiation: Analyzing platform power in the news industry. *Digital Journalism*, 11(8), 1391–1409.
<https://doi.org/10.1080/21670811.2022.2103011>
- Poell, T., Nieborg, D., & van Dijck, J. (2019). Platformisation. *Internet Policy Review*, 8(4), 1–13. <https://policyreview.info/concepts/platformisation>
- Pollicino, O., & De Gregorio, G. (2022). Constitutional democracy, platform powers and digital populism. *Constitutional Studies*, 8(1), 11–34.
<https://constitutionalstudies.wisc.edu/index.php/cs/article/view/87>
- Polyák, G. (2019). Media in Hungary: Three pillars of an illiberal democracy. In E. Połńska, & C. Beckett (Eds.), *Public service broadcasting and media systems in troubled European democracies* (pp. 279–303). Palgrave Macmillan.
https://doi.org/10.1007/978-3-030-02710-0_13
- Popiel, P. (2018). The tech lobby: Tracing the contours of new media elite lobbying power. *Communication, Culture and Critique*, 11(4), 566–585. <https://doi.org/10.1093/ccctcy027>
- Porter, E., & Wood, T. J. (2021). The global effectiveness of fact-checking: Evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom. *PNAS: Proceedings of the National Academy of Sciences*, 118(37), 1–7.
<https://doi.org/10.1073/pnas.2104235118>
- Posada, A., & Chen, G. (2018). Inequality in knowledge production: The integration of academic infrastructure by big publishers. *22nd International Conference on Electronic Publishing*. <https://pdfs.semanticscholar.org/4bb0/2d560d5d7b74560b9f9a6a6fe1493150f393.pdf>
- Posetti, J., & Bontcheva, K. (2020). *Disinfodemic: Deciphering COVID-19 disinformation*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000374416>
- Posetti, J., Shabbir, N., Douglas, O., & Gardiner, B. (2022). *The chilling: A global study of online violence against women journalists*. ICFJ (International Center for Journalists), UNESCO.
<https://www.icfj.org/our-work/chilling-global-study-online-violence-against-women-journalists>
- Pospieszna, P., Lown, P., & Dietrich, S. (2023). Building active youth in post-Soviet countries through civic education programmes: Evidence from Poland. *East European Politics*, 39(2), 321–344. <https://doi.org/10.1080/21599165.2022.2102608>

- Post, R. (2009). Hate speech. In I. Hare, & J. Weinstein (Eds.), *Extreme speech and democracy* (pp. 123–138). Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199548781.001.0001>
- Powell, A. B. (2015). Open culture and innovation: Integrating knowledge across boundaries. *Media, Culture & Society*, 37(3), 376–393. <https://doi.org/10.1177/0163443714567169>
- Powell, A. B. (2021). *Undoing optimization: Civic action in smart cities*. Yale University Press.
- Pradhan, S. (2021, April 15). Combating online misinformation: Diverging approaches in the United States and European Union. *Columbia Journal of Transnational Law Bulletin*.
<https://www.jtl.columbia.edu/bulletin-blog/combating-online-misinfo>
- Prasad, R. (2018). Ascendant India, digital India: How net neutrality advocates defeated Facebook's free basics. *Media, Culture & Society*, 40(3), 415–431.
<https://doi.org/10.1177/0163443717736117>
- Prior, M. (2005). News vs. entertainment: How increasing media choice widens gaps in political knowledge and turnout. *American Journal of Political Science*, 49(3), 577–592.
<https://doi.org/10.1111/j.1540-5907.2005.00143.x>
- Prochazka, F., & Schweiger, W. (2019). How to measure generalized trust in news media? An adaptation and test of scales. *Communication Methods and Measures*, 13(1), 26–42.
<https://doi.org/10.1080/19312458.2018.1506021>
- Procter, R., Arana-Catania, M., He, Y., Liakata, M., Zubiaga, A., Kochkina, E., & Zhao, R. (2023). Some observations on fact-checking work with implications for computational support. *ArXiv*. <https://arxiv.org/pdf/2305.02224>
- Procter, R., Webb, H., Jirotko, M., Burnap, P., Housley, W., Edwards, A., & Williams, M. (2019). A study of cyber hate on Twitter with implications for social media governance strategies. *ArXiv*. <https://doi.org/10.48550/arXiv.1908.11732>
- Program for Public Consultation. (2022). *Three-in-four voters favor reinstating net neutrality*. School of Public Policy, University of Maryland.
<https://publicconsultation.org/united-states/three-in-four-voters-favor-reinstating-net-neutrality>
- PublicSpaces International. (2024). How we think about AI is largely dictated by big tech.
<https://english.publicspaces.net/2024/06/03/how-we-think-about-ai-is-largely-dictated-by-big-tech>
- Puddephatt, A. (2021). *Letting the sun shine in: Transparency and accountability in the digital age*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000377231>
- Puehringer, S., Rath, J., & Griesebner, T. (2021). *The political economy of academic publishing*. *PLOS ONE*, 16(6), 1–21.
<https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0253226&type=printable>
- Punia, S., Mohan, S., Kakkar, J. M., & Bhandari, V. (Eds.). (2022). *Emerging trends in data governance*. National Law University, Delhi Press.
- Puppis, M., Mansell, R., & Van den Bulck, H. (Eds.). (2024). *Handbook of media and communication governance*. Edward Elgar Publishing. <https://doi.org/10.4337/9781800887206>
- Purandare, U., & Parkar, K. (2020). 'Eyes and ears': Surveillance in the Indian smart city. In J. C. Augusto (Ed.), *Handbook of smart cities* (pp. 1–32). Springer.
https://doi.org/10.1007/978-3-030-15145-4_60-1
- Purtova, N., & van Maanen, G. (2024). Data as an economic good, data as a commons, and data governance. *Law, Innovation and Technology*, 16(1), 1–42.
<https://doi.org/10.1080/17579961.2023.2265270>
- Pyo, J. Y. (2022). Different stakes, different struggles, and different practices to survive: News organizations and the spectrum of platform dependency. *New Media & Society*, 26(8), 4572–4588. <https://doi.org/10.1177/14614448221123279>
- Quelle, D., & Bovet, A. (2024). The perils and promises of fact-checking with large language models. *Frontiers in Artificial Intelligence*, 7, 1–14.
<https://doi.org/10.3389/frai.2024.1341697>
- Quintais, J. P., De Gregorio, G., & Magalhães, J. C. (2023). How platforms govern users' copyright-protected content: Exploring the power of private ordering and its implications. *Computer Law & Security Review*, 48, 1–25. <https://doi.org/10.1016/j.clsr.2023.105792>
- Quintavalla, A., & Temperman, J. (Eds.). (2023). *Artificial intelligence and human rights*. Oxford University Press. <https://doi.org/10.1093/law/9780192882486.001.0001>

- Radsch, C. C. (2022). *Making big tech pay for the news they use*. Center for International Media Assistance.
<https://www.cima.ned.org/publication/making-big-tech-pay-for-the-news-they-use/>
- Radsch, C. C. (2023a, June 2). *Envisioning a healthy information ecosystem*. CDT (Center for Democracy and Technology).
<https://cdt.org/insights/from-our-fellows-envisioning-a-healthy-information-ecosystem>
- Radsch, C. C. (2023b). Platformization and media capture: A framework for regulatory analysis of media-related platform regulations. *UCLA Journal of Law & Technology*, 28(2), 175–223. <https://uclajolt.com/platformization-and-media-capture-a-framework-for-regulatory-analysis-of-media-related-platform-regulations>
- Radsch, C. C. (2023c, August 8). The value of news content to Google is way more than you think. *Tech Policy Press*.
<https://techpolicy.press/the-value-of-news-content-to-google-is-way-more-than-you-think>
- Radsch, C. C. (2023d). *URGENT: Understanding and responding to global emerging news threats*. Internews.
<https://internews.org/wp-content/uploads/2023/03/29Mar-URGENT-Report-Final.pdf>
- Radsch, C. C. (2023e). *What makes for a healthy information ecosystem? New visual tool*.
<https://taicollaborative.org/what-makes-for-a-healthy-information-ecosystem-new-visual-tool>
- Radsch, C. C. (2024, August 22). The must-carry solution for the media's Google problem. *Washington Monthly*. <https://washingtonmonthly.com/2024/08/22/the-must-carry-solution-for-the-medias-google-problem>
- Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. *ACM Conference on Fairness, Accountability, and Transparency (FAT*)* 2020, 469–481. <http://dx.doi.org/10.2139/ssrn.3408010>
- Ragnedda, M., & Gladkova, A. (Eds.). (2020). *Digital inequalities in the Global South*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-32706-4>
- Ragnedda, M., & Ruiui, M. L. (2020). *Digital capital: A Bourdieusian perspective on the digital divide*. Emerald. <https://doi.org/10.1108/9781839095504>
- Rajkumar, M., & Ashraf, M. M. (2023). *Response to draft amendment to the IT (intermediary guidelines and digital media ethics code) rules, 2021*. IT for Change, India.
<https://itforchange.net/response-to-draft-amendment-to-it-intermediary-guidelines-and-digital-media-ethics-code-rules-2021>
- Ramiro, A., & Cruz, L. (2023). The grey-zones of public-private surveillance: Policy tendencies of facial recognition for public security in Brazilian cities. *Internet Policy Review*, 12(1), 1–28. <https://policyreview.info/articles/analysis/grey-zones-public-private-surveillance-policy-tendencies-facial-recognition>
- Ranaivoson, H., Afilipoaie, A., & Domazetovikj, N. (2022). *Media pluralism in the EU: A prospective look at the European media freedom act* [Policy Brief No. 64]. Vrije Universiteit Brussel, Studies in Media Innovation Technology Research Group.
https://www.researchgate.net/publication/369799716_Policy_Brief_64_Media_pluralism_in_the_EU_A_prospective_look_at_the_European_Media_Freedom_Act
- Rankin, J. (2023, February 2). MEPs launch site for EU officials to report 'shady lobbying' by big tech. *The Guardian*. <https://amp.theguardian.com/world/2023/feb/02/meeps-launch-site-for-eu-officials-to-report-shady-lobbying>
- Rantanen, T. (2024). *Dead men's propaganda: Ideology and utopia in comparative communications studies*. LSE Press. <https://press.lse.ac.uk/books/m/10.31389/lsepress.wmf>
- Ratner, Y., Dvir Gvirsman, S., & Ben-David, A. (2023). "Saving journalism from Facebook's death grip"? The implications of content-recommendation platforms on publishers and their audiences. *Digital Journalism*, 11(8), 1410–1431.
<https://doi.org/10.1080/21670811.2023.2180403>
- Recabarren, R., Carburnar, B., Hernandez, N., & Shafin, A. A. (2023, 2023). Strategies and vulnerabilities of participants in Venezuelan influence operations. *32nd USENIX Security Symposium*. www.usenix.org/conference/usenixsecurity23/presentation/recabarren
- Recuero, R. (2024). The platformization of violence: Toward a concept of discursive toxicity on social media. *Social Media + Society*, 10(1), 1–9. <https://doi.org/10.1177/20563051231224264>

- Recuero, R., Soares, F. B., & Gruzd, A. (2020). Hyperpartisanship, disinformation and political conversations on Twitter: The Brazilian presidential election of 2018. *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 569–578). <https://doi.org/10.1609/icwsm.v14i1.7324>
- Recuero, R., Soares, F. B., Vinhas, O., Volcan, T., Hüttner, L. R. G., & Silva, V. (2022). Bolsonaro and the far right: How disinformation about COVID-19 circulates on Facebook in Brazil. *International Journal of Communication*, 16(2022), 148–171. <https://ijoc.org/index.php/ijoc/article/viewFile/17724/3631>
- Reddi, M., Kuo, R., & Kreiss, D. (2023). Identity propaganda: Racial narratives and disinformation. *New Media & Society*, 25(8), 2201–2218. <https://doi.org/10.1177/14614448211029293>
- Reinemann, C., Stanyer, J., Scherr, S., & Legnante, G. (2012). Hard and soft news: A review of concepts, operationalizations and key findings. *Journalism*, 13(2), 221–239. <https://doi.org/10.1177/1464884911427803>
- Reis, O., Eneh, N. E., Ehimuan, B., Anyanwu, A., Olorunsogo, T., & Abrahams, T. O. (2024). Privacy law challenges in the digital age: A global review of legislation and enforcement. *International Journal of Applied Research in Social Sciences*, 6(1), 73–88. <https://doi.org/10.51594/ijarss.v6i1.733>
- Reisach, U. (2021). The responsibility of social media in times of societal and political manipulation. *European Journal of Operational Research*, 291(3), 906–917. <https://doi.org/10.1016/j.ejor.2020.09.020>
- Reiter, F., & Matthes, J. (2023). Correctives of the mainstream media? A panel study on mainstream media use, alternative digital media use, and the erosion of political interest as well as political knowledge. *Digital Journalism*, 11(5), 813–832. <https://doi.org/10.1080/21670811.2021.1974916>
- Renaldi, A. (2024, July 23). *In Indonesia, social media is a “hunting ground” for religious minorities*. Rest of World Reporting Global Tech Stories. <https://restofworld.org/2024/indonesia-muslim-influencers>
- Repnikova, M. (2018, September 6). China’s lessons for fighting fake news. *Foreign Policy*. <https://foreignpolicy.com/2018/09/06/chinas-lessons-for-fighting-fake-news>
- Repucci, S., & Slipowitz, A. (2022). *The global expansion of authoritarian rule*. Freedom House. <https://freedomhouse.org/report/freedom-world/2022/global-expansion-authoritarian-rule>
- Repucci, S., & Slipowitz, A. House. (2022). *Freedom in the world 2022: The global expansion of authoritarian rule*. Freedom House. <https://freedomhouse.org/report/freedom-world/2022/global-expansion-authoritarian-rule>
- Reuters. (2019, November 29). *China seeks to root out fake news and deepfakes with new online content rules*. <https://www.reuters.com/article/idUSKBN1Y30VT>
- Rexhepi, R. (2023). Content moderation: How the EU and the US approach is striking a balance between protecting free speech and protecting public interest. *Trento Student Law Review*, 5(1), 67–97. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4433201
- Reyna, A. (2024). DMA and DSA effective enforcement – Key to success. *Journal of Antitrust Enforcement*, 12(2), 320–324. <https://doi.org/10.1093/jaenfo/jnae018>
- RIA (Research ICT Africa). (2023a). *Data for democracy action: Countering disinformation in African elections through open access to data and journalism as a public good* [Workshop Outcome Report]. <https://researchictafrica.net/2023/12/11/data-for-democracy-action-countering-disinformation-in-african-elections-through-open-access-to-data-and-journalism-as-a-public-good/>
- RIA (Research ICT Africa). (2023b). *Information integrity on digital platforms* [RIA input on United Nations Policy Brief 8]. <https://researchictafrica.net/research/ria-input-on-united-nations-policy-brief-8-information-integrity-on-digital-platforms/>
- Ricaurte, P. (2022). Ethics for the majority world: AI and the question of violence at scale. *Media, Culture & Society*, 44(4), 726–745. <https://doi.org/10.1177/01634437221099612>
- Richards, N., & Hartzog, W. (2019). The pathologies of digital consent. *Washington University Law Review*, 96, 1461–1503. https://openscholarship.wustl.edu/law_lawreview/vol96/iss6/11/
- Richter, A. (2019). *Disinformation in the media under Russian law*. European Audiovisual Observatory. <https://rm.coe.int/disinformation-in-the-media-under-russian-law/1680967369>

- Richter, A. (2024). International legal responses to “propaganda for war” in modern warfare. *Journal of International Media & Entertainment Law*, 10(1), 54–80.
<https://www.swlaw.edu/curriculum/law-review-journals/journal-international-media-entertainment-law>
- Rieder, B., & Sire, G. (2014). Conflicts of interest and incentives to bias: A microeconomic critique of Google’s tangled position on the Web. *New Media & Society*, 16(2), 195–211.
<https://doi.org/10.1177/1461444813481195>
- Rieder, B., & Skop, Y. (2021). The fabrics of machine moderation: Studying the technical, normative, and organizational structure of Perspective API. *Big Data & Society*, 8(2), 1–16.
<https://doi.org/10.1177/20539517211046181>
- Riedl, M. J., Whipple, K. N., & Wallace, R. (2022). Antecedents of support for social media content moderation and platform regulation: The role of presumed effects on self and others. *Information, Communication & Society*, 25(11), 1632–1649.
<https://doi.org/10.1080/1369118X.2021.1874040>
- Ritzer, G., Ryan, J. M., Hayes, S., Elliot, M., & Jandrić, P. (2024). McDonaldization and artificial intelligence. *Postdigital Science and Education*, 6, 1320–1333.
<https://doi.org/10.1007/s42438-024-00475-3>
- Robb, G., & Hawthorne, R. (2019). Net neutrality and market power: The case of South Africa. *Telecommunications Policy*, 43(9), 1–15. <https://doi.org/10.1016/j.telpol.2019.03.003>
- Robbins, J. W. (2020, September 23). *Countering Russian disinformation* [Blog post]. CSIS (Center for Strategic and International Studies).
<https://www.csis.org/blogs/post-soviet-post/countering-russian-disinformation>
- Roberts, H., Hine, E., Taddeo, M., & Floridi, L. (2024). Global AI governance: Barriers and pathways forward. *International Affairs*, 100(3), 1275–1286.
<https://doi.org/10.1093/ia/iaae073>
- Roberts, T., & Ali, A. M. (2021). *Opening and closing online civic space in Africa: An introduction to the ten digital rights landscape reports*. Institute for Development Studies.
<https://hdl.handle.net/20.500.12413/15964>
- Roberts, T., & Bosch, T. (2023a). Introduction: Spaces of digital citizenship in Africa. In T. Roberts, & T. Bosch (Eds.), *Digital citizenship in Africa: Technologies of agency and repression* (pp. 1–32). Zed Books. <https://doi.org/10.5040/9781350324497>
- Roberts, T., & Bosch, T. (Eds.). (2023b). *Digital citizenship in Africa: Technologies of agency and repression*. Zed Books. <https://doi.org/10.5040/9781350324497>
- Roberts, T., Gitahi, J., Allam, P., Oboh, L., Oladapo, O., Appiah-Adjei, G., Galal, A., Kainja, J., Phiri, S., Abraham, K., Klovig Skelton, S., & Sheombar, A. (2023). *Mapping the supply of surveillance technologies to Africa: Case studies from Nigeria, Ghana, Morocco, Malawi, and Zambia*. Institute for Development Studies. <https://hdl.handle.net/20.500.12413/18120>
- Robertson, C. E., del Rosario, K. S., & Van Bavel, J. J. (2024). Inside the funhouse mirror factory: How social media distorts perceptions of norms. *Current Opinion in Psychology*, 60, 1–7. <https://doi.org/10.1016/j.copsyc.2024.101918>
- Robinson, L., Schulz, J., Dodel, M., Correa, T., Villanueva-Mansilla, E., Leal, S., Magallanes-Blanco, C., Rodriguez-Medina, L., Dunn, H. S., Levine, L., McMahon, R., & Khilnani, A. (2020). Digital inclusion across the Americas and Caribbean. *Social Inclusion*, 8(2), 244–259. <https://doi.org/10.17645/si.v8i2.2632>
- Roche, C., Wall, P. J., & Lewis, D. (2023). Ethics and diversity in artificial intelligence policies, strategies and initiatives. *AI and Ethics*, 3(4), 1095–1115.
<https://doi.org/10.1007/s43681-022-00218-9>
- Rodríguez-Ferrándiz, R. (2023). An overview of the fake news phenomenon: From untruth-driven to post-truth-driven approaches. *Media and Communication*, 11(2), 15–29.
<https://doi.org/10.17645/mac.v11i2.6315>
- Rogers, R. (2018). Digital traces in context: Otherwise engaged: Social media from vanity metrics to critical analytics. *International Journal of Communication*, 12, 450–472.
<https://ijoc.org/index.php/ijoc/article/view/6407>
- Rogers, R. (2020). The scale of Facebook’s problem depends upon how ‘fake news’ is classified. *Harvard Kennedy School Misinformation Review*, 1(6), 1–14.
<https://doi.org/10.37016/mr-2020-43>

- Rohn, U., von Rimscha, M. B., & Raats, T. (Eds.). (2024). *De Gruyter handbook of media economics*. De Gruyter. <https://doi.org/10.1515/9783110793444>
- Romano, A. (2013). *Politics and the press in Indonesia: Understanding an evolving political culture*. Routledge. <https://doi.org/10.4324/9780203037607>
- Romeu, A. (2023). *New digital law tightens clampdown on press freedom in Cuba*. Reporters Without Borders. <https://rsf.org/en/new-digital-law-tightens-clampdown-press-freedom-cuba>
- Ronzhyn, A., Cardenal, A. S., & Batlle Rubio, A. (2023). Defining affordances in social media research: A literature review. *New Media & Society*, 25(11), 3165–3188. <https://doi.org/10.1177/14614448221135187>
- Rosenberg, D. (2013). Data before the fact. In L. Gitelman (Ed.), “*Raw data*” is an oxymoron (pp. 15–40). MIT Press. <https://doi.org/10.7551/mitpress/9302.001.0001>
- Rosol, M., & Blue, G. (2022). From the smart city to urban justice in a digital age. *City*, 26(4), 684–705. <https://doi.org/10.1080/13604813.2022.2079881>
- Ross Arguedas, A., & Simon, F. M. (2023). *Automating democracy: Generative AI, journalism, and the future of democracy*. Balliol College, Oxford Internet Institute and Institute for Ethics in AI. <https://ora.ox.ac.uk/objects/uuid:0965ad50-b55b-4591-8c3b-7be0c587d5e7>
- Ross Arguedas, A., Robertson, C., Fletcher, R., & Nielsen, R. (2022a). *Echo chambers, filter bubbles, and polarisation: A literature review*. Reuters Institute for the Study of Journalism, University of Oxford. <https://ora.ox.ac.uk/objects/uuid:6e357e97-7b16-450a-a827-a92c93729a08>
- Ross Arguedas, A., Badrinathan, S., Mont’Alverne, C., Toff, B., Fletcher, R., & Nielsen, R. K. (2022b). “It’s a battle you are never going to win”: Perspectives from journalists in four countries on how digital media platforms undermine trust in news. *Journalism Studies*, 23(14), 1821–1840. <https://doi.org/10.1080/1461670X.2022.2112908>
- Ross Arguedas, A., Badrinathan, S., Mont’Alverne, C., Toff, B., Fletcher, R., & Nielsen, R. (2022c). *Snap judgements: How audiences who lack trust in news navigate information on digital platforms*. Reuters Institute for the Study of Journalism, University of Oxford. <https://ora.ox.ac.uk/objects/uuid:0cc7a92c-dc7c-41dd-81fa-33ca63ee0050>
- Roth, E. (2022, August 16). Google results for abortion clinics are misleading and politically fraught. *The Verge*. <https://www.theverge.com/2022/8/16/23307850/google-maps-results-abortion-clinics-crisis-pregnancy-centers>
- Rouvroy, A., & Berns, T. (2013). Algorithmic governmentality and prospects of emancipation. *Rezeaux*, 177(1), 163–196. <https://shs.cairn.info/journal-rezeaux-2013-1-page-163?lang=en>
- Roy, A. (2004, October 28). *Event summary: Poverty and the right to know*. Brookings Institution Commentary. <https://www.brookings.edu/articles/event-summary-poverty-and-the-right-to-know>
- Roy, R. (2024, March 23). *Big relief! Supreme court stays notification constituting fact-check unit!* Internet Freedom Foundation. <https://internetfreedom.in/sc-stays-notification-constituting-fcu>
- RSF (Reporters Without Borders). (2022). *Methodology used for compiling world press freedom index*. <https://rsf.org/en/index-methodologie-2022>
- RSF (Reporters Without Borders). (2024). *World press freedom index 2024 – journalism under political pressure*. <https://rsf.org/en/2024-world-press-freedom-index-journalism-under-political-pressure>
- RSF (Reporters Without Borders). (2025). *World press freedom index 2025: Economic fragility a leading threat to press freedom*. <https://rsf.org/en/rsf-world-press-freedom-index-2025-economic-fragility-leading-threat-press-freedom>
- Ruffo, G., Semeraro, A., Giachanou, A., & Rosso, P. (2023). Studying fake news spreading, polarisation dynamics, and manipulation by bots: A tale of networks and language. *Computer Science Review*, 47, 1–26. <https://doi.org/10.1016/j.cosrev.2022.100531>
- Ruggie, J. G. (2011). *Report of the special representative of the secretary-general on the issue of human rights and transnational corporations and other business enterprises: Guiding principles on business and human rights: Implementing the United Nations “protect, respect and remedy” framework*. United Nations General Assembly. <https://digitallibrary.un.org/record/705860?v=pdf>
- Rugh, W. A. (2000). *Arab mass media: Newspapers, radio, and television in Arab Politics*. Praeger Publishers.

- Ruohonen, J. (2003). David and Goliath: Privacy lobbying in the European Union. In J. Greenwood (Ed.), *The challenge of change in EU business associations* (pp. 139–149). Palgrave Macmillan. <https://doi.org/10.1057/9780230523234>
- Rusbridger, A. (2024, August 2). Elon Musk's misinformation machine made the horrors of Southport much worse. *Prospect*. <https://www.prospectmagazine.co.uk/ideas/media/67463/elon-musks-misinformation-machine-made-the-horrors-of-southport-much-worse>
- Ryan-Mosley, T. (2023, October 4). How generative AI is boosting the spread of disinformation and propaganda. *MIT Technology Review*. <https://www.technologyreview.com/2023/10/04/1080801/generative-ai-boosting-disinformation-and-propaganda-freedom-house/>
- Ryng, J., Guicherd, G., Saman, J. A., Choudhury, P., & Kellett, A. (2022). Internet shutdowns: A human rights issue. *The RUSI Journal*, 167(4-5), 50–63. <https://doi.org/10.1080/03071847.2022.2156234>
- Sacasas, L. M. (2021, August 12). The materiality of digital culture: It's not as invisible – or as disembodiment – as our devices lead us to believe. *Comment*. <https://comment.org/the-materiality-of-digital-culture>
- Sádaba, C., Salaverriá, R., & Bingué, X. (2023). Overcoming the age barrier: Improving older adults' detection of political disinformation with media literacy. *Media and Communication*, 11(4), 113–123. <http://dx.doi.org/10.17645/mac.v11i4.7090>
- Sadeghi, M., Dimitriadis, D., Arvanitis, L., Padovese, V., Pozzi, G., Badilini, S., Vercellone, C., Wang, M., Huet, N., Fishman, Z., Pfaller, L., Adams, N., & Wollen, M. (2025, May 5). Tracking AI-enabled misinformation: Over 1200 'unreliable AI-generated news' websites (and counting). *NewsGuard*. <https://www.newsguardtech.com/special-reports/ai-tracking-center>
- Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, 6(1), 1–12. <https://doi.org/10.1177/2053951718820549>
- Salehi, N., Irani, L. C., Bernstein, M. S., Alkhatib, A., Ogbe, E., Milland, K., & Clickhappier. (2015, April 18–23). We are dynamo: Overcoming stalling and friction in collective action for crowd workers. *CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 1621–1630. <http://dx.doi.org/10.1145/2702123.2702508>
- Samoilenko, S. A., & Suvorova, I. (2023). Artificial intelligence and deepfakes in strategic deception campaigns: The US and Russian experiences. In E. Pashentsev (Ed.), *The Palgrave handbook of malicious use of AI and psychological security* (pp. 507–529). Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-22552-9>
- Samoili, S., Lopez Cobo, M., Gomez Gutierrez, E., De Prato, G., Martinez-Plumed, F., & Delipetrev, B. (2020). AI WATCH: *Defining artificial intelligence*. European Commission. <https://publications.jrc.ec.europa.eu/repository/handle/JRC118163>
- Sampath, P. G. (2021). Governing artificial intelligence in an age of inequality. *Global Policy*, 12(S6), 21–31. <https://doi.org/10.1111/1758-5899.12940>
- Sančanin, B., & Penjišević, A. (2022). Use of artificial intelligence for the generation of media content. *Social Informatics Journal*, 1(1), 1–7. <https://doi.org/10.58898/sij.v1i1.01-07>
- Sanders, N. E., & Schneider, B. (2024, March 13). Let's not make the same mistakes with AI that we made with social media. *MIT Technology Review*. <https://www.technologyreview.com/2024/03/13/1089729/lets-not-make-the-same-mistakes-with-ai-that-we-made-with-social-media>
- Sankaran, V. (2024, February 22). Elon Musk's X takes on Indian government in freedom of speech row. *The Independent*. <https://www.independent.co.uk/tech/x-twitter-accounts-withheld-india-modi-b2500486.html>
- Santa Clara Principles.org. (2022). *The Santa Clara principles on transparency and accountability in content moderation*. <https://santaclaraprinciples.org>
- Santini, R. M., Agostini, L., Barros, C. E., Carvalho, D., Centeno de Rezende, R., Salles, D. G., Seto, K., Terra, C., & Tucci, G. (2018). Software power as soft power. A literature review on computational propaganda effects in public opinion and political process. *The Open Journal of Sociopolitical Studies*, 11(2), 1–29. <https://doi.org/10.1285/i20356609v11i2p332>

- Santini, R. M., Salles, D., & Barros, C. E. (2022). We love to hate George Soros: A cross-platform analysis of the Globalism conspiracy theory campaign in Brazil. *Convergence: The International Journal of Research into New Media Technologies*, 28(4), 983–1006. <https://doi.org/10.1177/13548565221085833>
- Santini, R. M., Salles, D., Borges, M., Medeiros, P., Traiano, H., Ferreira, F., Grael, F., de Pinho Mello, D., Moreira, A., Santos, M. L., Barros, C. E., Mattos, B., Melo, B., Haddad, J. G., Murakami, L., Loiola, D., do Carmo, V., Magalhães, T., Dias, B., ... Bastos, M. (2023). *The permanent campaign against the press in Brazil: Analysis of the media ecosystem and multiplatform digital advertising*. NetLab – Laboratory of Internet and Social Network Studies, Federal University of Rio de Janeiro (UFRJ).
- Santini, R. M., Salles, D., Tucci, G., Ferreira, F., & Grael, F. (2020). Making up audience: Media bots and the falsification of the public sphere. *Communication Studies*, 71(3), 466–487. <https://doi.org/10.1080/10510974.2020.1735466>
- Santos, P., & Ndlovu, K. (2022). Towards a journalism for justice: A normative overture. *African Journalism Studies*, 43(1), 70–88. <https://doi.org/10.1080/23743670.2021.2009533>
- Saurwein, F., & Spencer-Smith, C. (2020). Combating disinformation on social media: Multilevel governance and distributed accountability in Europe. *Digital Journalism*, 8(6), 820–841. <https://doi.org/10.1080/21670811.2020.1765401>
- Savitz, E. J. (2019, June 6). Google buys data-analytics firm Looker amid privacy concerns. *Baron's*. <https://www.barrons.com/articles/google-buys-data-analytics-firm-looker-privacy-concerns-51559837661>
- Schaake, M., & Fukuyama, F. (Eds.). (2023). *Digital technologies in emerging countries*. Stanford Cyber Policy Center. <https://cyber.fsi.stanford.edu/publication/digital-technologies-emerging-countries>
- Schaetz, N., Gagrčin, E., Toth, R., & Emmer, M. (2023). Algorithm dependency in platformized news use. *New Media & Society*, 27(3), 1360–1377. <https://doi.org/10.1177/14614448231193093>
- Schertel Mendes, L. S., & Kira, B. (2023, December 21). *The road to regulation of artificial intelligence: The Brazilian experience*. Internet Policy Review. <https://policyreview.info/articles/news/road-regulation-artificial-intelligence-brazilian-experience/1737>
- Schippers, B. (2020). Artificial intelligence and democratic politics. *Political Insight*, 11(1), 32–35. <https://doi.org/10.1177/2041905820911746>
- Schirch, L. (2021). Digital information, conflict and democracy. In L. Schirch (Ed.), *Social media impacts on conflict and democracy: The tectonic shift* (pp. 21–42). Routledge. <https://doi.org/10.4324/97818003087649>
- Schlesinger, P. (2020). After the post-public sphere. *Media, Culture & Society*, 42(7-8), 1545–1563. <https://doi.org/10.1177/0163443720948003>
- Schlossberg, J., & Freedman, D. (2020). Opening the gates: Plurality regulation in a post digital world. *Journal of Digital Media & Policy*, 11(2), 115–132. https://doi.org/10.1386/jdmp_00017_1
- Schmarzo, B. (2023). *AI & data literacy: Empowering citizens of data science*. Packt Publishing.
- Schmitt, V., Villa-Arenas, L. F., Feldhus, N., Meyer, J., Spang, R. P., & Möller, S. (2024). The role of explainability in collaborative human-AI disinformation detection. *FAccT'24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 2157–2174. <https://doi.org/10.1145/3630106.3659031>
- Schneider, N. (2022). Governable spaces: A feminist agenda for platform policy. *Internet Policy Review*, 11(1), 1–19. <https://doi.org/10.14763/2022.1.1628>
- Schoon, A., Mabweazara, H. M., Bosch, T., & Dugmore, H. (2020). Decolonising digital media research methods: Positioning African digital experiences as epistemic sites of knowledge production. *African Journalism Studies*, 41(4), 1–15. <https://doi.org/10.1080/23743670.2020.1865645>
- Schot, E. (2020). *Income models of independent media in difficult contexts*. Free Press Unlimited.
- Schranz, M., Schneider, J., & Eisenegger, M. (2018). Media trust and media use. In K. Otto, & A. Köhler (Eds.), *Trust in media and journalism: Empirical perspectives on ethics, norms, impacts and populism in Europe* (pp. 73–91). Springer Fachmedien.
- Schreurs, K., Quan-Haase, A., & Martin, K. (2017). Problematising the digital literacy paradox in the context of older adults' ICT use: Aging, media discourse, and self-determination. *Canadian Journal of Communication*, 42(2), 359–377. <http://doi.org/10.22230/cjc.2017v42n2a3130>

- Schudson, M. (1978). *Discovering the news: A social history of American newspapers*. Basic Books.
- Schudson, M. (2022). What does “trust in the media” mean? *Daedalus*, 151(4), 144–160. https://doi.org/10.1162/daed_a_01948
- Schünemann, W. J. (2022). A threat to democracies? An overview of theoretical approaches and empirical measurements for studying the effects of disinformation. In M. D. Cavelti, & A. Wenger (Eds.), *Cyber security politics: Socio-technological transformations and political fragmentation* (pp. 32–47). Routledge. <https://doi.org/10.4324/9781003110224>
- Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). *Towards a standard for identifying and managing bias in artificial intelligence*. US National Institute of Standards and Technology. <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf>
- Scott, M., Bunce, M., & Wright, K. (2022). The influence of news coverage on humanitarian aid: The bureaucrats’ perspective. *Journalism Studies*, 23(2), 167–186. <https://doi.org/10.1080/1461670X.2021.2013129>
- Segura, M. S., & Waisbord, S. (2019). Between data capitalism and data citizenship. *Television & New Media*, 20(4), 412–419. <https://doi.org/10.1177/1527476419834519>
- Seipp, T. J., Fathaigh, R. O., & van Drunen, M. (2023a). Defining the ‘media’ in Europe: Pitfalls of the proposed European media freedom act. *Journal of Media Law*, 15(1), 39–51. <https://doi.org/10.1080/17577632.2023.2240998>
- Seipp, T. J., Helberger, N., de Vreese, C., & Ausloos, J. (2023b). Dealing with opinion power in the platform world: Why we really have to rethink media concentration law. *Digital Journalism*, 11(8), 1542–1567. <https://doi.org/10.1080/21670811.2022.2161924>
- Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019, January 29). Fairness and abstraction in sociotechnical systems. *FAT* ’19: Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68. <https://doi.org/10.1145/3287560.3287598>
- Sepas, A., Bangash, A. H., Alraoui, O., El Emam, K., & El-Hussuna, A. (2022). Algorithms to anonymize structured medical and healthcare data: A systematic review. *Frontiers in Bioinformatics*, 2, 1–11. <https://doi.org/10.3389/fbinf.2022.984807>
- Settle, J. E. (2018). *Frenemies: How social media polarizes America*. Cambridge University Press. <https://doi.org/10.1017/9781108560573>
- Sevignani, S., Theine, H., & Tröger, M. (2025). Toward media environment capture: A theoretical contribution on the influence of big tech on news media. *International Journal of Communication*, 19, 804–824. <https://ijoc.org/index.php/ijoc/article/view/21987>
- Sey, A., Shilongo, K., Hlomani, H., & Diouf, D. (2022). Sub-Saharan Africa. In H. Wasserman (Ed.), *Meeting the challenges of information disorder in the Global South* (pp. 138–189). IDRC Research Results. <https://idl-bnc-idrc.dspacedirect.org/bitstreams/aafc0871-9bfe-4d11-bd8e-cb7921064621/download>
- Shahbaz, A., Funk, A., & Vesteinsson, K. (2022). *Countering an authoritarian overhaul of the internet*. Freedom House. <https://freedomhouse.org/report/freedom-net/2022/countering-authoritarian-overhaul-internet>
- Shahin, S. (2019). Facing up to Facebook: How digital activism, independent regulation, and mass media foiled a neoliberal threat to net neutrality. *Information, Communication & Society*, 22(1), 1–17. <https://doi.org/10.1080/1369118X.2017.1340494>
- Shapero, J. (2024, August 30). Musk calls Brazilian judge ‘evil dictator’ amid X feud. *The Hill*. <https://thehill.com/policy/technology/4855584-musk-slams-brazilian-judge-threatens-x/>
- Shapiro, R. (2019). Transnational networks of literacy and materiality: Coltan, sexual violence, and digital literacy. *College English*, 82(2), 204–225. <https://doi.org/10.58680/ce201930626>
- Sheombar, A., & Skelton, S. K. (2023). Follow the surveillance: A breadcrumb trail of surveillance technology exports to Africa. In M. R. Jones, A. S. Mukherjee, D. Thapa, & Y. Zheng (Eds.), *After Latour: Globalisation, inequity and climate change* (pp. 241–261). IFIP WG 8.2 & 9.4 Working Conference Proceedings, Springer Nature. <https://link.springer.com/book/10.1007/978-3-031-50154-8>
- Sherman, J. (2022, September 5). India’s sudden reversal on privacy will affect the global internet. *Slate*. <https://slate.com/technology/2022/09/india-data-protection-bill-fourth-way.html>

- Shin, D.-H., & Lee, M.-K. (2017). Public value mapping of network neutrality: Public values and net neutrality in Korea. *Telecommunications Policy*, 41(3), 208–224. <https://doi.org/10.1016/j.telpol.2016.12.012>
- Shin, D.-H., Kee, K. F., & Shin, E. Y. (2022). Algorithm awareness: Why user awareness is critical for personal privacy in the adoption of algorithmic platforms? *International Journal of Information Management*, 65, 1–13. <https://doi.org/10.1016/j.ijinfomgt.2022.102494>
- Shirikov, A. (2024). Fake news for all: How citizens discern disinformation in autocracies. *Political Communication*, 41(1), 45–65. <https://doi.org/10.1080/10584609.2023.2257618>
- Shleifer, A. (2004). *Does competition destroy ethical behavior?* [NBER Working paper 10269]. National Bureau of Economic Research. <https://www.nber.org/papers/w10269>
- Shoemaker, P. J., & Vos, T. (2009). *Gatekeeping theory*. Routledge. <https://doi.org/10.4324/9780203931653>
- Siapera, E. (2023). The challenge of far-right media. In S. Farrell, E. Siapera, & G. Souvlis (Eds.), *Radical journalism: Resurgence, reform, reaction* (pp. 135–149). Routledge.
- Silverstone, R. (2007). *Media and morality: On the rise of the mediapolis*. Polity Press.
- Simon, F. M. (2022). Uneasy bedfellows: AI in the news, platform companies and the issue of journalistic autonomy. *Digital Journalism*, 10(10), 1832–1854. <https://doi.org/10.1080/21670811.2022.2063150>
- Simon, F. M. (2024). *Artificial intelligence in the news: How AI retools, rationalizes, and reshapes journalism and the public arena*. Tow Center for Journalism, Oxford Internet Institute, University of Oxford. <https://doi.org/10.7916/ncm5-3v06>
- Simon, F. M., & Isaza-Ibarra, L. F. (2023). *AI in the news: Reshaping the information ecosystem?* Balliol College, Torch, Reuters Institute, University of Oxford. <https://ora.ox.ac.uk/objects/uuid:9947240c-06d3-42c2-9a23-57ff4559b63c/files/sj3860857t>
- Simon, F. M., Altay, S., & Mercier, H. (2023). Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown. *Harvard Kennedy School (HKS) Misinformation Review*, 4(5), 1–11. <https://doi.org/10.37016/mr-2020-127>
- Singapore Statutes. (2019). *Protection from online falsehoods and manipulation act 2019*. Government of Singapore. <https://tinyurl.com/cs2jafue>
- Singh, P. J. (2020). *Economic rights in a data-based society: Collective data ownership, workers' rights, and the role of the public sector*. Public Services International, Friedrich Ebert Foundation; IT for Change, India. <https://tinyurl.com/472u6y9f>
- Singh, P. J., & Gurumurthy, A. (2021). *Economic governance of data: Balancing individualist-property approaches with a community rights framework*. IT for Change, India. <https://itforchange.net/sites/default/files/1880/Economic-governance-of-data.pdf>
- Sinpeng, A. (2020). Digital media, political authoritarianism, and internet controls in Southeast Asia. *Media, Culture & Society*, 42(1), 25–39. <https://doi.org/10.1177/0163443719884052>
- Sirbu, A., Pedreschi, D., Giannotti, F., & Kertész, J. (2019). Algorithmic bias amplifies opinion fragmentation and polarization: A bounded confidence model. *PLOS ONE*, 14(3), 1–20. <https://doi.org/10.1371/journal.pone.0213246>
- Skalli, L. H. (2023). Disruptive social entrepreneurship from Bahrain: The case of Esra'a Al Shafei. In L. H. Skalli, & N. Eltantawy (Eds.), *The Palgrave handbook of gender, media and communication in the Middle East and North Africa* (pp. 567–583). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-11980-4_31
- Skibinski, M. (2021). Special report: Top brands are sending \$2.6 billion to misinformation websites each year. *NewsGuard*. <https://www.newsguardtech.com/special-reports/brands-send-billions-to-misinformation-websites-newsguard-comscore-report>
- Skovsgaard, M., & Andersen, K. (2020). Conceptualizing news avoidance: Towards a shared understanding of different causes and potential solutions. *Journalism Studies*, 21(4), 459–476. <https://doi.org/10.1080/1461670X.2019.1686410>
- Sleibi, A. (2023). Wartime opportunities: The sway of Russian disinformation on the MENA region. In *Wartime-Russia yearbook 2023* (pp. 312–317). EIMed (European Institute of the Mediterranean). <https://www.iemed.org/med-yearbook/iemed-mediterranean-yearbook-2023>
- Smith, J. R. (2019, January 29). *IBM Research releases 'diversity in faces' dataset to advance study of fairness in facial recognition systems*. Phys.org. <https://phys.org/news/2019-01-ibm-diversity-dataset-advance-fairness.html>

- Söderström, O., & Datta, A. (Eds.). (2023). *Data power in action: Urban data politics in times of crisis*. Bristol University Press.
- Sodré, M. (2021). *A sociedade incivil: Mídia, iliberalismo e finanças* [The uncivil society: media, illiberalism and finance]. Editora Vozes.
- Solomon, S., Polataiko, M., & Hayes, H. A. (2021). Platform responsibility and regulation in Canada: Considerations on transparency, legislative clarity, and design. *Harvard Journal of Law & Technology*, 34, 1–18. <https://jolt.law.harvard.edu/digest/platform-responsibility-and-regulation-in-canada-considerations-on-transparency-legislative-clarity-and-design>
- Solove, D. J., & Hartzog, W. (2024, July 25). The great scrape: The clash between scraping and privacy. *California Law Review*. Advance online publication. <https://doi.org/10.2139/ssrn.4884485>
- Souter, D. (2024, October 1). *Inside the digital society: Governing AI*. APC (Association for Progressive Communications). <https://www.apc.org/en/blog/inside-digital-society-governing-ai>
- South Africa Government. (2024). *National policy on data and cloud*. Communications & Digital Technologies Department, Republic of South Africa. https://www.gov.za/sites/default/files/gcis_document/202406/50741gen2533.pdf
- Spitale, G., Biller-Andorno, N., & Germani, F. (2023). AI model GPT-3 (dis)informs us better than humans. *Science Advances*, 9(26), 1–9. <https://doi.org/10.1126/sciadv.adh1850>
- Splichal, S. (2022a). *Datafication of public opinion and the public sphere: How extraction replaced expression of opinion*. Anthem Press. <https://www.jstor.org/stable/j.ctv2s2pp3n>
- Splichal, S. (2022b). The public sphere in the twilight zone of publicness. *European Journal of Communication*, 37(2), 198–215. <https://doi.org/10.1177/02673231211061490>
- SSD EFF (Surveillance Self-Defense, Electronic Frontier Foundation). (2023). *Surveillance self-defense*. <https://ssd.eff.org>
- Stamboliev, E. (2023). Proposing a postcritical AI literacy: Why we should worry less about algorithmic transparency and more about citizen empowerment. *Media Theory*, 7(1), 202–232. <https://hal.science/hal-04221538>
- Stanly, M. (2024, July 4). *The future of work: Ensuring AI literacy for all*. INDIAai. <https://indiaai.gov.in/article/the-future-of-work-ensuring-ai-literacy-for-all>
- Star, S. L., & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research*, 7(1), 111–134. <https://www.jstor.org/stable/23010792>
- Steenen, S., Grøndahl Larsen, A. M., Hågvær, Y. B., & Fonn, B. K. (2019). What does digital journalism studies look like? *Digital Journalism*, 7(3), 320–342. <https://doi.org/10.1080/21670811.2019.1581071>
- Stefaniya, A. P., & Pierson, J. (2023). Algorithmic governmentality, digital sovereignty, and agency affordances: Extending the possible fields of action. *Weizenbaum Journal of the Digital Society*, 3(2), 1–30. <https://doi.org/10.34669/WI.WJDS/3.2.2>
- Štětka, V., & Mihelj, S. (2024a). The rise of the illiberal public sphere. In V. Štětka, & S. Mihelj (Eds.), *The illiberal public sphere: Media in polarized societies* (pp. 25–59). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-54489-7_2
- Štětka, V., & Mihelj, S. (2024b). *The illiberal public sphere: Media in polarized societies*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-54489-7>
- Stevenson, C. (2016, December 6). *Seattle surveillance ordinance memo*. Technology Law and Public Policy Clinic. <https://digitalcommons.law.uw.edu/cgi/viewcontent.cgi?article=1002&context=techclinic>
- Stiefel, L., Currie, M., Musiani, F., Sandoz, A., Silvast, A., & Williams, R. (2024). Preface: Governance by infrastructure. *First Monday*, 29(2), 1–4. <https://firstmonday.org/ojs/index.php/fm/article/view/13559/11406>
- Stigler Committee. (2019). *Stigler committee on digital platforms: Final report*. Stigler Center for the Study of the Economy and the State, University of Chicago Booth School of Business. <https://research.chicagobooth.edu/stigler/media/news/committee-on-digital-platforms-final-report>
- Stolton, S., & Maksimov, V. (2020, March 30). *Orbán to rule by decree with new powers to 'silence critics'*. Euractiv. <https://www.euractiv.com/section/global-europe/news/orban-to-rule-by-decree-with-new-powers-to-silence-critics>

- Stray, J., Iyer, R., & Larrauri, H. P. (2023). *The algorithmic management of polarization and violence on social media*. Knight First Amendment Institute, Columbia University.
<http://knightcolumbia.org/content/the-algorithmic-management-of-polarization-and-violence-on-social-media>
- Strömbäck, J., & Shehata, A. (2019). The reciprocal effects between political interest and TV news revisited: Evidence from four panel surveys. *Journalism & Mass Communication Quarterly*, 96(2), 473–496. <https://doi.org/10.1177/1077699018793998>
- Strömbäck, J., Falasca, K., & Kruikemeier, S. (2018). The mix of media use matters: Investigating the effects of individual news repertoires on offline and online political participation. *Political Communication*, 35(3), 413–432.
<https://doi.org/10.1080/10584609.2017.1385549>
- Strömbäck, J., Tsifti, Y., Boomgaarden, H., Damstra, A., Lindgren, E., Vliegthart, R., & Lindholm, T. (2020). News media trust and its impact on media use: Toward a framework for future research. *Annals of the International Communication Association*, 44(2), 139–156. <https://doi.org/10.1080/23808985.2020.1755338>
- Stroud, N. J. (2017). Selective exposure theories. In K. Kenski, & K. H. Jamieson (Eds.), *The Oxford handbook of political communication* (pp. 531–548). Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199793471.001.0001>
- Stroud, N. J., & Lee, J. K. (2013). Perceptions of cable news credibility. *Mass Communication and Society*, 16(1), 67–88. <https://doi.org/10.1080/15205436.2011.646449>
- Strover, S., Choi, J., & Schrubbe, A. (2024). Broadband, rural contexts and local economic dynamics. *Telecommunications Policy*, 48(3), 1–12.
<https://doi.org/10.1016/j.telpol.2024.102720>
- Suber, P. (2012). *Open access*. MIT Press.
- Suchman, L. (2023). The uncontroversial ‘thingness’ of AI. *Big Data & Society*, 10(2), 1–5.
<https://doi.org/10.1177/20539517231206794>
- Suiter, J., & Fletcher, R. (2020). Polarization and partisanship: Key drivers of distrust in media old and new? *European Journal of Communication*, 35(5), 484–501.
<https://doi.org/10.1177/0267323120903685>
- Sultan, M., Tump, A. N., Ehmann, N., Lorenz-Spreen, P., et al. (2024). Susceptibility to online misinformation: A systematic meta-analysis of demographic and psychological factors. *PsyArXiv*. <https://doi.org/10.31234/osf.io/7u4fg>
- Sunstein, C. (2007). *Republic.com 2.0*. Princeton University Press.
- Szakacs, J., & Bogner, É. (2021). *The impact of disinformation campaigns about migrants and minority groups in the EU*. European Parliament, Directorate General for External Policies of the Union.
[https://www.europarl.europa.eu/thinktank/en/document/EXPO_IDA\(2021\)653641](https://www.europarl.europa.eu/thinktank/en/document/EXPO_IDA(2021)653641)
- Szczepkowski, D., & Szczepkowski, P. (2021). Adware as one form of disinformation. *Studia Administracji i Bezpieczeństwa*, 1(10), 197–208.
<https://www.ceol.com/search/article-detail?id=1004222>
- Talamanca, G., & Arfini, S. (2022). Through the newsfeed glass: Rethinking filter bubbles and echo chambers. *Philosophy and Technology*, 35(1), 1–34.
<https://doi.org/10.1007/s13347-021-00494-z>
- Tambini, D. (2021a). *Media freedom*. Polity.
- Tambini, D. (2021b). What is journalism? The paradox of media privilege. *European Human Rights Law Review*, 5, 523–539. https://eprints.lse.ac.uk/112681/1/What_is_journalism_The_paradox_of_media_privilege_clean.pdf
- Tan, C. (2020). *Regulating content on social media: Copyright, terms of service and technological features*. Creative Media Partners.
- Tandoc, E. C. Jr., & Vos, T. P. (2016). The journalist is marketing the news: Social media in the gatekeeping process. *Journalism Practice*, 10(8), 950–966.
<https://doi.org/10.1080/17512786.2015.1087811>
- Taneja, H., Webster, J. G., Malthouse, E. C., & Ksiazek, T. B. (2012). Media consumption across platforms: Identifying user-defined repertoires. *New Media & Society*, 14(6), 951–968.
<https://doi.org/10.1177/1461444811436146>

- Tang, S., Willnat, L., & Zhang, H. (2021). Fake news, information overload, and the third-person effect in China. *Global Media and China*, 6(4), 492–507.
<https://doi.org/10.1177/20594364211047369>
- Tapper, J. (2024, August 3). The far right has moved online, where its voice is more dangerous than ever. *The Observer*. <http://www.theguardian.com/politics/article/2024/aug/03/the-far-right-has-moved-online-where-its-voice-is-more-dangerous-than-ever>
- Tarrant, A., & Cowen, T. (2022). Big tech lobbying in the EU. *The Political Quarterly*, 93(2), 218–226. <https://doi.org/10.1111/1467-923X.13127>
- Tate-Mosely, R. (2023, June 20). The movement to limit face recognition tech might finally get a win. *MIT Technology Review*.
<http://www.technologyreview.com/2023/07/20/1076539/face-recognition-massachusetts-test-police>
- Tay, L. Q., Lewandowsky, S., Hurlstone, M. J., Kurz, T., & Ecker, U. K. H. (2024). Thinking clearly about misinformation. *Communications Psychology*, 2(1), 1–5.
<https://doi.org/10.1038/s44271-023-00054-5>
- Taylor, L., Mukiri-Smith, H., Petročnik, T., Savolainen, L., & Martin, A. (2022). (Re)making data markets: An exploration of the regulatory challenges. *Law, Innovation and Technology*, 14(2), 355–394. <https://doi.org/10.1080/17579961.2022.2113671>
- Teachout, Z., & Khan, L. (2014). Market structure and political law: A taxonomy of power. *Duke Journal of Constitutional Law & Public Policy*, 9, 37–74.
<https://doi.org/10.2139/ssrn.2490525>
- Tech Hive Advisory Center for Law & Innovation. (2024). *State of AI regulation in Africa: Trends and developments report*. <http://www.techhiveadvisory.africa/report/state-of-ai-regulation-in-africa-trends-and-developments>
- Teixeira da Silva, J. A. (2021). How to shape academic freedom in the digital age? Are the retractions of opinionated papers a prelude to “cancel culture” in academia? *Current Research in Behavioral Sciences*, 2, 1–6. <https://doi.org/10.1016/j.crbeha.2021.100035>
- Tenove, C. (2020). Protecting democracy from disinformation: Normative threats and policy responses. *The International Journal of Press/Politics*, 25(3), 517–537.
<https://doi.org/10.1177/1940161220918740>
- Teyssou, D., Leung, J.-M., Apostolidis, E., Apostolidis, K., Papadopoulos, S., Zampoglou, M., Papadopoulou, O., & Mezaris, V. (2017). The InVID plug-in: Web video verification on the browser. *Proceedings of the First International Workshop on Multimedia Verification* (pp. 23–30). ACM (Association for Computing Machinery).
<https://doi.org/10.1145/3132384.3132387>
- Thakur, D., & Madrigal, D. H. (2022). *An unrepresentative democracy: How disinformation and online abuse hinder women of color political candidates in the United States*. Center for Democracy & Technology.
<https://cdt.org/insights/an-unrepresentative-democracy-how-disinformation-and-online-abuse-hinder-women-of-color-political-candidates-in-the-united-states/>
- The Majority Report w/ Sam Seder. (2023). *Data cartels and our data* [Video]. YouTube.
<https://www.youtube.com/watch?v=E-VWLIFUxIA>
- The Tierra Común Network. (2023). *Resisting data colonialism: A practical intervention*. Institute of Network Cultures, Amsterdam University. https://networkcultures.org/wp-content/uploads/2023/12/ResistingDataColonialism_INC2023_TOD50.pdf
- The White House. (2025a). *Removing barriers to American leadership in artificial intelligence*. <https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/>
- The White House. (2025b). *Fact sheet: President Donald J. Trump ends the taxpayer subsidization of biased media*. <https://www.whitehouse.gov/fact-sheets/2025/05/fact-sheet-president-donald-j-trump-ends-the-taxpayer-subsidization-of-biased-media/>
- Thomas, E. (2022). *Conspiracy clickbait: This one weird trick will undermine democracy*. ISD Institute for Strategic Studies.
https://www.isdglobal.org/digital_dispatches/conspiracy-clickbait/
- Thompson, A., & Kemp, S. (2024, January 31). *Digital 2024: 5 billion social media users*. We Are Social. <https://wearesocial.com/uk/blog/2024/01/digital-2024-5-billion-social-media-users>
- Thussu, D. K. (2022). De-colonizing global news-flows: A historical perspective. *Journalism Studies*, 23(13), 1578–1592. <https://doi.org/10.1080/1461670X.2022.2083007>

- Timcke, S., & Hlomani, H. (2024). *How might AI reshape democracy on the African continent?* Research ICT Africa. <https://researchictafrica.net/2024/02/05/how-might-ai-reshape-democracy-on-the-african-continent/>
- Timcke, S., Makumbirofa, S., Rens, A., Omar, N., & Hlomani, H. (2024). *The political economy of African AI: A primer on concepts, contexts, considerations and capitalism.* Research ICT Africa. <https://researchictafrica.net/research/a-new-primer-on-the-political-economy-of-african-ai/>
- Toff, B., & Kalogeropoulos, A. (2020). All the news that's fit to ignore: How the information environment does and does not shape news avoidance. *Public Opinion Quarterly*, 84(S1), 366–390. <https://doi.org/10.1093/poq/nfaa016>
- Toff, B., & Nielsen, R. K. (2018). “I just Google it”: Folk theories of distributed discovery. *Journal of Communication*, 68(3), 636–657. <https://doi.org/10.1093/joc/jqy009>
- Toff, B., & Palmer, R. (2019). Explaining the gender gap in news avoidance: “News-is-for-men” perceptions and the burdens of caretaking. *Journalism Studies*, 20(11), 1563–1579. <https://doi.org/10.1080/1461670X.2018.1528882>
- Toff, B., Badrinathan, S., Mont’Alverne, C., Ross Arguedas, A., Fletcher, R., & Nielsen, R. (2021a). *Overcoming indifference: What attitudes towards news tell us about building trust.* Reuters Institute for the Study of Journalism, University of Oxford. <https://ora.ox.ac.uk/objects/uuid:ae2a4f80-62a3-45a5-947b-9ef638a8dc64>
- Toff, B., Badrinathan, S., Mont’Alverne, C., Ross Arguedas, A., Fletcher, R., & Nielsen, R. (2021b). *Listening to what trust in news means to users: Qualitative evidence from four countries.* Reuters Institute for the Study of Journalism, University of Oxford. <https://ora.ox.ac.uk/objects/uuid:f11227fa-45fd-48c9-80e5-64807691f80e>
- Toff, B., Palmer, R., & Nielsen, R. K. (2023). *Avoiding the news: Reluctant audiences for journalism.* Columbia University Press.
- Tollefson, J. (2024, June 26). ‘Vindicated’: Embattled misinformation researchers celebrate key US Supreme Court decision. *Nature*. <https://www.nature.com/articles/d41586-024-01766-2>
- Tomaz, T., & Trappel, J. (2022). News media in European democracies and beyond: Stable structural conditions but notorious deficits. *De Europa*, 5(2), 17–34. <https://doi.org/10.13135/2611-853X/7200>
- Tong, C., Gill, H., Li, J., Valenzuela, S., & Rojas, H. (2020). “Fake news is anything they say!” – Conceptualization and weaponization of fake news among the American public. *Mass Communication & Society*, 23(5), 755–778. <https://doi.org/10.1080/15205436.2020.1789661>
- Tonnies, F. (1957). *Community & society* [*Gemeinschaft und Gesellschaft*] [C. P. Loomis, Trans.]. Harper & Row (Original work published 1887)
- Trappel, J. (Ed.). (2019). *Digital media inequalities: Policies against divides, distrust and discrimination.* Nordicom, University of Gothenburg. <https://urn.kb.se/resolve?urn=urn:nbn:se:norden:org:diva-5550>
- Trauthig, I. (2024). *Diaspora communities and computational propaganda on messaging apps.* CIGI (Centre for International Governance Innovation). <https://www.cigionline.org/publications/diaspora-communities-and-computational-propaganda-on-messaging-apps/>
- Tsanni, A. (2024, January 19). How satellite images and AI could help fight spatial apartheid in South Africa. *MIT Technology Review*. <https://www.technologyreview.com/2024/01/19/1086837/satellite-images-ai-spatial-apartheid-south-africa>
- Tsfati, Y., & Ariely, G. (2014). Individual and contextual correlates of trust in media across 44 countries. *Communication Research*, 41(6), 760–782. <https://doi.org/10.1177/0093650213485972>
- Tsfati, Y., Boomgaarden, H. G., Strömbäck, J., Vliegenthart, R., Damstra, A., & Lindgren, E. (2020). Causes and consequences of mainstream media dissemination of fake news: Literature review and synthesis. *Annals of the International Communication Association*, 44(2), 157–173. <https://doi.org/10.1080/23808985.2020.1759443>
- Tuchman, G. (1972). Objectivity as strategic ritual: An examination of newsmen’s notions of objectivity. *American Journal of Sociology*, 77(4), 660–679. <https://doi.org/10.1086/225193>
- Tucker, E. (2023, June 14). Our future inside the fifth column – or, what chatbots are really for. *Tech Policy Press*. <https://www.techpolicy.press/our-future-inside-the-fifth-column-or-what-chatbots-are-really-for>

- Tucker, J. A., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). *Social media, political polarization, and political disinformation: A review of the literature*. Hewlett Foundation.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3144139
- Tully, M. (2022). Everyday news use and misinformation in Kenya. *Digital Journalism*, 10(1), 109–127. <https://doi.org/10.1080/21670811.2021.1912625>
- Tully, M., Madrid-Morales, D., Wasserman, H., Gondwe, G., & Ireri, K. (2022). Who is responsible for stopping the spread of misinformation? Examining audience perceptions of responsibilities and responses in six sub-Saharan African countries. *Digital Journalism*, 10(5), 679–697. <https://doi.org/10.1080/21670811.2021.1965491>
- Turillazzi, A., Taddeo, M., Floridi, L., & Casolari, F. (2023). The Digital Services Act: An analysis of its ethical, legal, and social implications. *Law, Innovation and Technology*, 15(1), 83–106. <https://doi.org/10.1080/17579961.2023.2184136>
- Turow, J., Lelkes, Y., Draper, N., & Waldman, A. E. (2023). *Americans can't consent to companies' use of their data: They admit they don't understand it, say they're helpless to control it, and believe they're harmed when firms use their data – making what companies do illegitimate*. Annenberg School for Communications, University of Pennsylvania.
https://www.asc.upenn.edu/sites/default/files/2023-02/Americans_Cant_Consent.pdf
- Tusikov, N. (2021). Internet platforms weaponizing chokepoints. In D. W. Drezner, H. Farrell, & A. Newman (Eds.), *The uses and abuses of weaponized interdependence* (pp. 133–148). Brookings Institution Press.
- Tuters, M., & Hagen, S. (2020). (((They))) rule: Memetic antagonism and nebulous othering on 4chan. *New Media & Society*, 22(12), 2218–2237.
<https://doi.org/10.1177/1461444819888746>
- Twomey, J., Ching, D., Aylett, M. P., Quayle, M., Linehan, C., & Murphy, G. (2023). Do deepfake videos undermine our epistemic trust? A thematic analysis of tweets that discuss deepfakes in the Russian invasion of Ukraine. *PLOS ONE*, 18(10), 1–22.
<https://doi.org/10.1371/journal.pone.0291668>
- Tyagi, K., Sanders, A. K., & Cauffman, C. (Eds.). (2024). *Digital platforms, competition law, and regulation: Comparative perspectives*. Bloomsbury Publishing.
- UAE (United Arab Emirates). (2018). *UAE National Strategy for Artificial Intelligence*. National Program for Artificial Intelligence. <https://ai.gov.ae/wp-content/uploads/2021/07/UAE-National-Strategy-for-Artificial-Intelligence-2031.pdf>
- UK DSIT (Department for Science, Innovation & Technology). (2023). *The Bletchley declaration by countries attending the AI safety summit, 1-2 November 2023* UK Dept for Science, Innovation & Technology.
<https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>
- UK Government. (2003). *Communications Act 2003*.
<https://www.legislation.gov.uk/ukpga/2003/21/contents>
- UK Government. (2023). *Online Safety Act 2023*. <https://www.legislation.gov.uk/ukpga/2023/50>
- Ulbricht, L., & Egbert, S. (2024). In Palantir we trust? Regulation of data analysis platforms in public security. *Big Data & Society*, 11(3), 1–15. <https://doi.org/10.1177/20539517241255108>
- UN. (1948). *Universal Declaration of Human Rights*.
<https://www.un.org/en/about-us/universal-declaration-of-human-rights>
- UN. (1966). *International Covenant on Civil and Political Rights* [General Assembly resolution 2200A (XXI)]. <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>
- UN. (2019, May). *United Nations strategy and plan of action on hate speech*. United Nations Secretary-General. https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_EN.pdf
- UN. (2020). *Promoting digital public goods*. United Nations Office of the Secretary General's Envoy on Technology. https://www.un.org/digital-emerging-technologies/sites/www.un.org/techenvoy/files/general/Digital_Public_Goods_Summary_PDF.pdf
- UN. (2022). *Countering disinformation for the promotion and protection of human rights and fundamental freedoms report of the secretary-general*. United Nations General Assembly.
<https://digitallibrary.un.org/record/3987886?v=pdf>

- UN. (2023a). *Information integrity on digital platforms* [Our common agenda policy brief 8]. <https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-information-integrity-en.pdf>
- UN. (2023b). *Interim report: Governing AI for humanity*. United Nations AI Advisory Board. https://www.un.org/sites/un2.un.org/files/ai_advisory_body_interim_report.pdf
- UN. (2023c). *UN Plan of action on the safety of journalists and the issue of impunity*. https://www.ohchr.org/sites/default/files/documents/issues/journalists/2023-01-31/un-plan-on-safety-journalists_en.pdf
- UN. (2024a). *Governing AI for humanity: Final report*. United Nations, AI Advisory Board. https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf
- UN. (2024b). *Pact for the future, global digital compact, and declaration on future generations*. https://www.un.org/sites/un2.un.org/files/sotf-pact_for_the_future_adopted.pdf
- UN. (2024c). *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development*. General Assembly. <https://digitallibrary.un.org/record/4043244/?v=pdf>
- UN. (2024d). *United Nations global principles for information integrity: Recommendations for multi-stakeholder action*. <https://www.un.org/sites/un2.un.org/files/un-global-principles-for-information-integrity-en.pdf>
- UN. (2024e). *What is hate speech? Understanding hate speech*. <https://www.un.org/en/hate-speech/understanding-hate-speech/what-is-hate-speech>
- UNESCO. (2022a). *Global standards for media and information literacy, curricula development guidelines*. <https://tinyurl.com/2nv52am2>
- UNESCO. (2022b). *K-12 AI curricula: A mapping of government-endorsed AI curricula*. <https://unesdoc.unesco.org/ark:/48223/pf0000380602>
- UNESCO. (2022c). *Recommendation on the ethics of artificial intelligence*. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- UNESCO. (2022d). *World trends in freedom of expression and media development: Global report 2021/22*. <https://www.unesco.org/en/world-media-trends>
- UNESCO. (2023a). *Addressing hate speech through education: A guide for policy makers*. <https://unesdoc.unesco.org/ark:/48223/pf0000384872>
- UNESCO. (2023b). *Guidelines for the governance of digital platforms: Safeguarding freedom of expression and access to information through a multi-stakeholder approach*. <https://unesdoc.unesco.org/ark:/48223/pf0000387339>
- UNESCO. (2023c). *World press freedom day 2023, shaping a future of rights: Freedom of expression as a driver for all other human rights; draft concept note*. <https://unesdoc.unesco.org/ark:/48223/pf0000384177.locale=en>
- UNESCO. (2024). *Observatory of killed journalists*. <https://www.unesco.org/en/safety-journalists/observatory>
- UNESCO. (n.d.). *Media and information literacy*. <https://www.unesco.org/en/media-information-literacy>
- Urman, A., & Makhortykh, M. (2025). The silence of the LLMs: Cross-lingual analysis of guardrail-related political bias and false information prevalence in ChatGPT, Google Bard (Gemini), and Bing Chat. *Telematics and Informatics*, 96, 1–16. <https://doi.org/10.1016/j.tele.2024.102211>
- US Congress. (1996a). *Health Insurance Portability and Accountability Act (HIPAA)*. UPub. L. 104–191. <https://www.govinfo.gov/content/pkg/PLAW-104publ191/pdf/PLAW-104publ191.pdf>
- US Congress. (1996b). *Telecommunications Act of 1996, section 230 protection for private blocking and screening of offensive material*. Pub. LA. 104–104. <https://tinyurl.com/ybn684ra>
- US Congress. (2013). *Children's Online Privacy Protection Act 1998*. U15 USC 6501–6505, Federal Trade Commission Amendment. <https://www.ftc.gov/legal-library/browse/rules/childrens-online-privacy-protection-rule-coppa>
- US Congress. (2025). *Tools to Address known exploitation by immobilizing technological deepfakes on websites and networks act or the take it down act*. US Congress, Public Law No. 119–12. <https://www.congress.gov/bill/119th-congress/senate-bill/146>

- US District Court. (2024, August 5). *US v. Google*. United States District Court for the District of Columbia, Case No. 20-cv-3010 (APM). <https://www.texasattorneygeneral.gov/sites/default/files/images/press/Google%20Search%20Engine%20Monopoly%20Ruling.pdf>
- US Government. (2020). *National artificial intelligence initiative act of 2020*. US Government H. R. 6395-1136, Division E. <https://www.aip.org/sites/default/files/aipcorp/images/fyi/pdf/national-ai-initiative-act-final.pdf>
- US Register of Copyrights. (2024). *Copyright and artificial intelligence: Part 1: Digital replicas*. United States Copyright Office. <https://www.copyright.gov/newsnet/2025/1060.html>
- US Register of Copyrights. (2025). *Copyright and artificial intelligence: Part 2: Copyrightability*. <https://www.copyright.gov/newsnet/2025/1060.html>
- US State of California. (2018). *California consumer privacy Act (CCPA)*. State of California Department of Justice. <https://oag.ca.gov/privacy/ccpa>
- US-China Commission. (2022). *China's evolving data governance regime*. US-China Economic and Security Review Commission. <https://www.uscc.gov/research/chinas-evolving-data-governance-regime>
- Usher, N. (2023). The real problems with the problem of news deserts: Toward rooting place, precision, and positionality in scholarship on local news and democracy. *Political Communication*, 40(2), 238–253. <https://doi.org/10.1080/10584609.2023.2175399>
- Uyheng, J., & Carley, K. M. (2024). Online hate in the Philippines: The role of narratives and networks. *International Review of Law, Computers & Technology*, 38(2), 230–244. <https://doi.org/10.1080/13600869.2023.2295104>
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13. <https://doi.org/10.1177/2056305120903408>
- Vaidhyanathan, S. (2011). *The googlization of everything: How one company is transforming culture, commerce, and community – and why we should worry*. University of California Press.
- Vaidhyanathan, S. (2022). *Anti-social media: How Facebook disconnects us and undermines democracy*. Oxford University Press. <https://doi.org/10.1093/oso/9780190056544.001.0001>
- Valenzuela, S., Bachmann, I., Lawrence, R. G., & Gil De Zúñiga, H. (2023). Politics and media. *Journalism & Mass Communication Quarterly*, 100(4), 808–825. <https://doi.org/10.1177/10776990231203542>
- Valenzuela, S., Halpern, D., & Araneda, F. (2022). A downward spiral? A panel study of misinformation and media trust in Chile. *The International Journal of Press/Politics*, 27(2), 353–373. <https://doi.org/10.1177/19401612211025238>
- Valenzuela, S., Muñoz, C., & Santos, M. (2024). Social media and belief in misinformation in Mexico: A case of maximal panic, minimal effects? *The International Journal of Press/Politics*, 29(3), 646–666. <https://doi.org/10.1177/19401612221088988>
- van der Linden, S. (2023). *Foolproof: Why we fall for misinformation and how to build immunity*. Fourth Estate.
- Van der Spuy, A. (2023). Unveiling digital development risks: The uncertain outcomes of promoting digital technologies in the guise of development [Doctoral dissertation, London School of Economics and Political Science, England]. <https://etheses.lse.ac.uk/4662/>
- van der Vlist, F., Helmond, A., & Ferrari, F. (2024). Big AI: Cloud infrastructure dependence and the industrialisation of artificial intelligence. *Big Data & Society*, 11(1), 1–16. <https://doi.org/10.1177/20539517241232630>
- van Dijk, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197–208. <https://doi.org/10.24908/ss.v12i2.4776>
- van Dijk, J. (2020). Seeing the forest for the trees: Visualizing platformization and its governance. *New Media & Society*, 23(9), 2801–2819. <https://doi.org/10.1177/1461444820940293>
- van Dijk, J., & Poell, T. (2013). Understanding social media logic. *Media and Communication*, 1(1), 2–14. <https://doi.org/10.12924/mac2013.01010002>
- van Dijk, J., de Waal, M., & Poell, T. (2018a). *The platform society: Public values in a connective world*. Oxford University Press. <https://doi.org/10.1093/oso/9780190889760.001.0001>

- van Dijk, J., Poell, T., & de Waal, M. (2018b). News. In J. van Dijk, T. Poell, & M. de Waal, *The platform society: Public values in a connective world*. (pp. 49–72). Oxford University Press. <https://doi.org/10.1093/oso/9780190889760.001.0001>
- van Drunen, M. Z., & Noroozian, A. (2024). How to design data access for researchers: A legal and software development perspective. *Computer Law & Security Review*, 52, 1–15. <https://doi.org/10.1016/j.clsr.2024.105946>
- Van Natta, J., Masadeh, S., & Hamilton, B. (2023, April 16). Investigating the impacts of YouTube's content policies on journalism and political discourse. *Proceedings of the ACM on Human-Computer Interaction*, 7, CSCW1. <https://doi.org/10.1145/3579499>
- Vanoli, C. (2024, April 19). Moving AI governance from principles to practice. *ITU News*. <https://www.itu.int/hub/2024/04/moving-ai-governance-from-principles-to-practice>
- Vasudevan, C., Linehan, M., Byers, C., Brooks, N. N., & Freeman, L. (2024). *Responsible generative AI: An examination of ongoing efforts to tame this powerful technology*. OMG and IBM. https://www.objectmanagementgroup.org/wp-content/uploads/sites/8/2024/03/JOI_20240318_Responsible_Generative_AI_standalone.pdf
- Vayadande, K., Singh, V., Sultanpure, K., Deshpande, D., Patil, H., Patil, A., Pathak, A., Pareek, S., & Patil, D. (2024). Empowering data sovereignty: Decentralized image sharing through blockchain and InterPlanetary File System. *International Journal of Intelligent Systems and Applications in Engineering*, 12(14S), 223–235. <https://ijisae.org/index.php/IJISAE/article/view/4660>
- Vázquez-Herrero, J., Negreira-Rey, M.-C., & López-García, X. (2022a). Let's dance the news! How the news media are adapting to the logic of TikTok. *Journalism*, 23(8), 1717–1735. <https://doi.org/10.1177/1464884920969092>
- Vázquez-Herrero, J., Negreira-Rey, M.-C., & Sixto-García, J. (2022b). Mind the gap! Journalism on social media and news consumption among young audiences. *International Journal of Communication*, 16, 21. <https://ijoc.org/index.php/ijoc/article/view/19643>
- Veale, M., Matus, K., & Gorwa, R. (2023). AI and global governance: Modalities, rationales, tensions. *Annual Review of Law and Social Science*, 19, 255–275. <https://doi.org/10.1146/annurev-lawsocsci-020223-040749>
- V-Dem Institute. (2025). *Democracy report 2025: 25 years of autocratization—democracy trumped?* https://v-dem.net/documents/60/V-dem-dr_2025_lowres.pdf
- Veiga, A. K., Saraiva, A. M., Chapman, A. D., Morris, P. J., Gendreau, C., Schigel, D., & Robertson, T. J. (2017). A conceptual framework for quality assessment and management of biodiversity data. *PLOS ONE*, 12(6), e0178731. <https://doi.org/10.1371/journal.pone.0178731>
- Veliz, C. (2021). *Privacy is power: Why and how you should take back control of your data*. Melville House.
- Velkova, J. (2024, February 20). *Dismantling public values, one data center at the time*. NordMedia Network. <https://nordmedianetwork.org/latest/news/dismantling-public-values-one-data-center-at-the-time>
- Ven, K., Verelst, J., & Mannaert, H. (2008). Should you adopt open source software? *IEEE Software*, 25(3), 54–59. <https://doi.org/10.1109/MS.2008.73>
- Verdegem, P. (2021, February 5). Tim Berners-Lee's plan to save the internet: Give us back control of our data. *The Conversation*. <https://theconversation.com/tim-berners-lees-plan-to-save-the-internet-give-us-back-control-of-our-data-154130>
- Verdegem, P. (2022). Dismantling AI capitalism: The commons as an alternative to the power concentration of big tech. *AI & Society*, 39, 727–737. <https://doi.org/10.1007/s00146-022-01437-8>
- Verhulst, S. G. (2023). Steering responsible AI: A case for algorithmic pluralism. *ArXiv*. <https://doi.org/10.48550/arXiv.2311.12010>
- Verhulst, S. G. (2024). Governing mediation in the data ecosystem: Lessons from media governance for overcoming data asymmetries. In M. Puppis, R. Mansell, & H. Van den Bulck (Eds.), *Handbook of media and communication governance* (pp. 372–383). Edward Elgar Publishing. <https://doi.org/10.4337/9781800887206>
- Verhulst, S. G., & Schüür, F. (2023, November 20). Interwoven realms: Data governance as the bedrock for AI governance [Blog post]. *Medium*. <https://medium.com/data-policy/interwoven-realms-data-governance-as-the-bedrock-for-ai-governance-ffd56a6a4543>

- Verma, N., Fleischmann, K. R., & Koltai, K. S. (2018). Demographic factors and trust in different news sources. *Proceedings of the Association for Information Science and Technology*, 55(1), 524–533. <https://doi.org/10.1002/pra2.2018.14505501057>
- Verza, S., Blagojev, T., Da Cots Leite Borges, D., Kermer, J. E., Trevisan, M., & Reviglio Della Venaria, U. (2024). *Uncovering news deserts in Europe: Risks and opportunities for local and community media in the EU*. European University Institute. <https://cadmus.eui.eu/handle/1814/76652>
- Viljoen, S. (2021). A relational theory of data governance. *Yale Law Journal*, 131(2), 573–654. <https://www.yalelawjournal.org/feature/a-relational-theory-of-data-governance>
- Villi, M., Aharoni, T., Tenenboim-Weinblatt, K., Boczkowski, P. J., Hayashi, K., Mitchelstein, E., Tanaka, A., & Kligler-Vilenchik, N. (2022). Taking a break from news: A five-nation study of news avoidance in the digital era. *Digital Journalism*, 10(1), 148–164. <https://doi.org/10.1080/21670811.2021.1904266>
- Vincent, R. C., & Nordenstreng, K. (Eds.). (2015). *Towards equity in global communication? MacBride update* (2nd ed.). Hampton Press.
- Vinhas, O., & Bastos, M. (2022). Fact-checking misinformation: Eight notes on consensus reality. *Journalism Studies*, 23(4), 448–468. <https://doi.org/10.1080/1461670X.2022.2031259>
- Voelkel, J. G., Stagnaro, M. N., Chu, J. Y., Pink, S. L., Mernyk, J. S., Redekopp, C., Ghezae, I., Cashman, M., Adjodah, D., Allen, L. G., Allis, L. V., Baleria, G., Ballantyne, N., Van Bavel, J. J., Blunden, H., Braley, A., Bryan, C. J., Celniker, J. B., Cikara, M., ... Willer, R. (2024). Megastudy testing 25 treatments to reduce antidemocratic attitudes and partisan animosity. *Science*, 386(290), 1–12. <https://doi.org/10.1126/science.adh4764>
- Vogels, E. A., & Anderson, M. (2019). *Americans and digital knowledge*. Pew Research Center. <https://www.pewresearch.org/internet/2019/10/09/americans-and-digital-knowledge/>
- von Ingersleben-Seip, N. (2023). Competition and cooperation in artificial intelligence standard setting: Explaining emergent patterns. *Review of Policy Research*, 40(5), 781–810. <https://doi.org/10.1111/ropr.12538>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Voss, W. G., & Pernot-Leplay, E. (2024). China data flows and power in the era of Chinese big tech. *Northwestern Journal of International Law and Business*, 44(1), 1–68. <https://scholarlycommons.law.northwestern.edu/njilb/vol44/iss1/1>
- VPUU (Violence Prevention thorough Urban Upgrading). (2019). *CitySpec: A mobile inspection tool improving service delivery*. <http://vpuu.org.za/safe-node-area/cityspec-service-delivery>
- Vuorikari, R., Kluzer, S., & Punie, Y. (2022). *DigComp 2.2: The digital competence framework for citizens – With new examples of knowledge, skills and attitudes*. European Commission, Joint Research Centre. <https://publications.jrc.ec.europa.eu/repository/handle/JRC128415>
- Wagner, B., Kettemann, M. C., Vieth-Dilman, K., & Montgomery, S. (Eds.). (2025). *Research handbook on human rights and digital technology: Global politics, law and international relations*. Edward Elgar Publishing. <https://doi.org/10.4337/9781785367724>
- Wahl-Jorgensen, K., & Carlson, M. (2021). Conjecturing fearful futures: Journalistic discourses on deepfakes. *Journalism Practice*, 15(6), 803–820. <https://doi.org/10.1080/17512786.2021.1908838>
- Waisbord, S. (2023). Mob censorship revisited: Questions, findings, and challenges. *Digital Journalism*, 11(10), 1761–1768. <https://doi.org/10.1080/21670811.2023.2214905>
- Wakefield, J. (2016, June 15). *Social media 'outstrips TV' as news source for young people*. BBC News. <https://www.bbc.com/news/uk-36528256>
- Waldman, A. E. (2020). Cognitive biases, dark patterns, and the 'privacy paradox'. *Current Opinion in Psychology*, 31, 105–109. <https://doi.org/10.1016/j.copsyc.2019.08.025>
- Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2020). Fact-checking: A meta-analysis of what works and for whom. *Political Communication*, 37(3), 350–375. <https://doi.org/10.1080/10584609.2019.1668894>
- WAN-IFRA (World Association of News Publishers). (2023). *Gauging generative AI's impact on newsrooms: Survey: Newsroom executives share their experiences so far*. World Association of News Publishers. <https://wan-ifra.org/insight/gauging-generative-ais-impact-in-newsrooms>

- Wang, B., Rau, P.-L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology*, 42(9), 1324–1337. <https://doi.org/10.1080/0144929X.2022.2072768>
- Wang, C., Boerman, S. C., Kroon, A. C., Möller, J., & de Vreese, C. H. (2024). The artificial intelligence divide: Who is the most vulnerable? *New Media & Society*, 1–23. Online First. <https://doi.org/10.1177/14614448241232345>
- Wang, W. W. (2023). China's digital transformation: Data-empowered state capitalism and social governmentality. *The African Journal of Information and Communication*, (31), 1–13. <https://doi.org/10.23962/ajic.i31.16296>
- Wardle, C. (2018). The need for smarter definitions and practical, timely empirical research on information disorder. *Digital Journalism*, 6(8), 951–963. <https://doi.org/10.1080/21670811.2018.1502047>
- Wardle, C. (2023). Misunderstanding misinformation. *Issues in Science and Technology*, 29(3), 38–40. <https://doi.org/10.58875/ZAUD1691>
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making* [Council of Europe Report DGO(2017)09]. <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>
- Warschauer, M. (2004). *Technology and social inclusion: Rethinking the digital divide*. MIT Press. <https://doi.org/10.7551/mitpress/6699.001.0001>
- Washington, J. (2023). *Combating misinformation and fake news: The potential of AI and media literacy education*. <http://dx.doi.org/10.2139/ssrn.4580385>
- Wasko, J., Murdock, G., & Sousa, H. (Eds.). (2011). *The handbook of political economy of communications*. John Wiley & Sons. <https://doi.org/10.1002/9781444395402>
- Wasserman, H. (2018). *Media, geopolitics, and power: A view from the Global South*. University of Illinois Press.
- Wasserman, H. (2020a). Fake news from Africa: Panics, politics and paradigms. *Journalism*, 21(1), 3–16. <https://doi.org/10.1177/1464884917746861>
- Wasserman, H. (2020b). The state of South African media: A space to contest democracy. *Publizistik*, 65(3), 451–465. <https://doi.org/10.1007/s11616-020-00594-4>
- Wasserman, H. (Ed.). (2022). *Meeting the challenges of information disorder in the Global South*. IDRC (International Development Research Centre, Canada). <https://idl-bnc-idrc.dspacedirect.org/bitstreams/aafc0871-9bfe-4d11-bd8e-eb7921064621/download>
- Wasserman, H., & Madrid-Morales, D. (2019). An exploratory study of “fake news” and media trust in Kenya, Nigeria and South Africa. *African Journalism Studies*, 40(1), 107–123. <https://doi.org/10.1080/23743670.2019.1627230>
- WEF (World Economic Forum). (2024). *Global risks report 2024 19th edition: Insight report*. <https://www.weforum.org/publications/global-risks-report-2024>
- WEF (World Economic Forum). (2025). *Global risks report 2025 20th edition: Insight report*. <https://www.weforum.org/publications/global-risks-report-2025/>
- WEF (World Economic Forum). (2024, April 25). How to manage AI's energy demand – today and in the future. <https://www.weforum.org/agenda/2024/04/how-to-manage-ais-energy-demand-today-tomorrow-and-in-the-future>
- Weismueller, J., Gruner, R. L., Harrigan, P., Coussement, K., & Wang, S. (2024). Information sharing and political polarisation on social media: The role of falsehood and partisanship. *Information Systems Journal*, 34(3), 854–893. <https://doi.org/10.1111/isj.12453>
- Werthner, H., Ghezzi, C., Kramer, J., Nida-Rümelin, J., Nuseibeh, B., Prem, E., & Stanger, A. (Eds.). (2024). *Introduction to digital humanism: A textbook*. Springer. <https://doi.org/10.1007/978-3-031-45304-5>
- West, E., & McAllister, M. P. (Eds.). (2023). *The Routledge companion to advertising and promotional culture*. Routledge. <https://doi.org/10.4324/9781003124870>
- West, S. M. (2019). Data capitalism: Redefining the logics of surveillance and privacy. *Business & Society*, 58(1), 20–41. <https://doi.org/10.1177/0007650317718185>
- Whittaker, M. (2021). The steep cost of capture. *Interactions*, 28(6), 50–55. <https://doi.org/10.1145/3488666>

- Wiggers, K. (2019, January 29). *IBM releases diversity in faces, a dataset of over 1 million annotations to help reduce facial recognition bias*. VentureBeat.
<https://venturebeat.com/ai/ibm-releases-diversity-in-faces-a-dataset-of-over-1-million-annotations-to-help-reduce-facial-recognition-bias>
- Wihbey, J. (2024). *AI and epistemic risk for democracy: A coming crisis of public knowledge?*
<https://doi.org/10.2139/ssrn.4805026>
- Wike, R., Fetterolf, J., Smerkovich, M., Austin, S., Gubbala, S., & Lippert, J. (2024, February 28). *Representative democracy remains a popular ideal, but people around the world are critical of how it's working* [Press release]. Pew Research Center.
https://www.pewresearch.org/global/wp-content/uploads/sites/2/2024/02/gap_2024.02.28_democracy-closed-end_report.pdf
- Wilding, D. (2021). Regulating news and disinformation on digital platforms: Self-regulation or prevarication? *Journal of Telecommunications and the Digital Economy*, 9(2), 11–46.
<https://doi.org/10.18080/jtde.v9n2.415>
- Willems, W. (2014a). Beyond normative dewesternization: Examining media culture from the vantage point of the Global South. *The Global South*, 8(1), 7–23.
<https://doi.org/10.2979/globalsouth.8.1.7>
- Willems, W. (2014b). Provincializing hegemonic histories of media and communication studies: Toward a genealogy of epistemic resistance in Africa. *Communication Theory*, 24(4), 415–434.
<https://doi.org/10.1111/comt.12043>
- Williams, A. (2023). *Surveillance nation: The real cost of Amazon's customer obsession*.
<https://peertube.dair-institute.org/w/oAWnan9PXTvS3CyutatQcg>
- Windwehr, S., & York, J. C. (2020, October 8). *Facebook's most recent transparency report demonstrates the pitfalls of automated content moderation*. Electronic Frontier Foundation.
<https://www.eff.org/deeplinks/2020/10/facebook-s-most-recent-transparency-report-demonstrates-pitfalls-automated-content>
- Wineburg, S., & McGrew, S. (2019). Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. *Teachers College Record*, 121(11), 1–40. <https://doi.org/10.1177/016146811912101102>
- Winseck, D. (2022). The broken internet and platform regulation: Promises and perils. In T. Flew, & F. R. Martin (Eds.), *Digital platform regulation: Global perspectives on internet governance* (pp. 229–257). Palgrave Macmillan.
https://doi.org/10.1007/978-3-030-95220-4_12
- Winseck, D., & Pooley, J. D. (2017). A reply to Faulhaber, Singer, and Urschel's curious tale of economics and common carriage (net neutrality) at the FCC. *International Journal of Communication*, 11, 2702–2733. <https://ijoc.org/index.php/ijoc/article/view/7543>
- Wintour, P. (2024, February 26). US leading global alliance to counter foreign government disinformation. *The Guardian*. <https://www.theguardian.com/technology/2024/feb/26/us-leading-global-alliance-to-counter-foreign-government-disinformation>
- WIPO (World Intellectual Property Organization). (2024). *Getting the innovation ecosystem ready for AI: An IP policy toolkit*. Frontier Technologies Division.
www.wipo.int/publications/en/details.jsp?id=4711
- Wojcowski, T. (2022). Conspiracy theories, right-wing populism and foreign policy: The case of the alternative for Germany. *Journal of International Relations and Development*, 25, 130–148. <https://doi.org/10.1057/s41268-021-00218-y>
- Wood, S. (2024). *Impact of regulation on children's digital lives*. Digital Futures for Children, London School of Economics and Political Science and 5Rights Foundation.
<https://www.digital-futures-for-children.net/our-work/regulation-impact>
- Wu, C.-J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, M., Maeng, K., Chang, G., Behram, F. A., Huang, J., Bai, C., Gschwind, M., Gupta, A., Ott, M., Melnikov, A., Candido, S., Brooks, D., Chauhan, G., Lee, B., Lee, H. S., ... Hazelwood, K. (2022). Sustainable AI: Environmental implications, challenges and opportunities. *ArXiv*.
<https://doi.org/10.48550/arXiv.2111.00364>
- Wu, J., Gan, W., Chen, Z., Wan, S., & Lin, H. (2023). AI-generated content (AIGC): A survey. *ArXiv*. <https://arxiv.org/pdf/2304.06632>
- Wu, T. (2003). Network neutrality, broadband discrimination. *Journal of Telecommunications and High Technology Law*, 2, 141–179. <https://doi.org/10.2139/ssrn.388863>

- Wu, T. (2017). *The attention merchants: The epic struggle to get inside our heads*. Atlantic Books.
- Wu, T. (2018). *The curse of bigness: Antitrust in the new gilded age*. Columbia Global Reports.
- Wulf, A. J., & Seizov, O. (2022). "Please understand we cannot provide further information": Evaluating content and transparency of GDPR-mandated AI disclosures. *AI & SOCIETY*, 39, 235–256. <https://doi.org/10.1007/s00146-022-01424-z>
- Wuyckens, G., Landry, N., & Fastrez, P. (2022). Untangling media literacy, information literacy, and digital literacy: A systematic meta-review of core concepts in media education. *Journal of Media Literacy Education*, 14(1), 168–182. <https://doi.org/10.23860/JMLE-2022-14-1-12>
- X. (2024). *About community notes on X*. X Help Center. <https://help.x.com/en/using-x/community-notes>
- Xu, J. (2023, April 25). What does China's newly launched national data bureau mean to China and global data governance? *Internet Policy Review*. <https://policyreview.info/articles/news/chinas-national-data-bureau-and-global-data-governance>
- Xu, Z., Jain, S., & Kankanhalli, M. (2024). Hallucination is inevitable: An innate limitation of large language models. *ArXiv*. <https://doi.org/10.48550/arXiv.2401.11817>
- Yan, B., Li, K., Xu, M., Dong, Y., Zhang, Y., Ren, Z., & Cheng, X. (2024). On protecting the data privacy of large language models (LLMs): A survey. *ArXiv*. <https://doi.org/10.48550/arXiv.2403.05156>
- Yan, S., & Yang, Y. (2021). Education Informatization 2.0 in China: Motivation, framework, and vision. *ECNU Review of Education*, 4(2), 410–428. <https://doi.org/10.1177/2096531120944929>
- Yang, Y., & Bennett, L. (2022). Interactive propaganda: How Fox News and Donald Trump co-produced false narratives about the COVID-19 crisis. In P. Van Aelst, & J. G. Blumler (Eds.), *Political communication in the time of coronavirus* (pp. 83–100). Routledge.
- Yates, S., & Carmi, E. (Eds.). (2024). *Digital inclusion: International policy and research*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-28930-9>
- Ye, Q. (2023). Comparison of the transparency of fact-checking: A global perspective. *Journalism Practice*, 17(10), 2263–2282. <https://doi.org/10.1080/17512786.2023.2211555>
- Yoo, C. S. (2024). Network slicing and net neutrality. *Telecommunications Policy*, 48(2), 1–10. <https://doi.org/10.1016/j.telpol.2023.102619>
- York, J. C. (2022, October 29). Elon Musk doesn't know what it takes to make a digital town square. *MIT Technology Review*. <https://www.technologyreview.com/2022/10/29/1062417/elon-musk-twitter-takeover-global-democracy-activists>
- Ytre-Arne, B., & Moe, H. (2021). Doomscrolling, monitoring and avoiding: News use in COVID-19 pandemic lockdown. *Journalism Studies*, 22(13), 1739–1775. <https://doi.org/10.1080/1461670X.2021.1952475>
- Zammit, M., Voulgari, I., & Yannakakis, G. N. (2021). *The road to AI literacy education: From pedagogical needs to tangible game design* [University of Malta paper for European Conference on Games Based Learning]. https://www.academia.edu/78830274/The_Road_to_AI_Literacy_Education_from_Pedagogical_Needs_to_Tangible_Game_Design
- Zelizer, B. (2019). Why journalism is about more than digital technology. *Digital Journalism*, 7(3), 343–350. <https://doi.org/10.1080/21670811.2019.1571932>
- Zhang, W., Ong, L. M., & Findlay, M. (2023). Digital self-determination: An alternative paradigm for emerging economies. In M. Findlay, L. M. Ong, & W. Zhang (Eds.), *Elgar companion to regulating AI and big data in emerging economies* (pp. 158–179). Edward Elgar Publishing. <https://doi.org/10.4337/9781785362408>
- Zhang, Y., Lukito, J., Su, M.-H., Suk, J., Xia, Y., Kim, S. J., Doroshenko, L., & Wells, C. (2021). Assembling the networks and audiences of disinformation: How successful Russian IRA Twitter accounts built their followings, 2015–2017. *Journal of Communication*, 71(2), 305–331. <https://doi.org/10.1093/joc/jqaa042>
- Zhao, Y., & Chen, J. (2022). A survey on differential privacy for unstructured data content. *ACM Computing Surveys*, 54(10s), Article 207, 1–28. <https://doi.org/10.1145/3490237>
- Zhu, J. (2022). AI ethics with Chinese characteristics? Concerns and preferred solutions in Chinese academia. *AI & Society*, 39(3), 1261–1274. <https://doi.org/10.1007/s00146-022-01578-w>

- Zhu, Q., Weeks, B. E., & Kwak, N. (2024). Implications of online incidental and selective exposure for political emotions: Affective polarization during elections. *New Media & Society*, 26(1), 450–472. <https://doi.org/10.1177/14614448211061336>
- Zibi Fama, P. A. (2024). *La démocratie à l'épreuve de la désinformation genrée en Afrique centrale et de l'Ouest* [Democracy put to the test by gendered disinformation in central and west Africa]. Organisation Internationale de la francophonie (ODIL). <https://odil.org/wp-content/uploads/2023/12/La-democratie-a-lepreuve-de-la-desinformation-genree%E2%80%AFen-Afrique-Centrale-et-de-lOuest.pdf>
- Znojek, B. (2020). The growing problem of disinformation in Latin America. *PISM Bulletin*. Polski Instytut Spraw Międzynarodowych (PISM). <https://www.cceol.com/search/gray-literature-detail?id=1167507>
- Zuazo, N., & Aruguete, N. (2021). ¿Polarización política o digital? un ecosistema con todos los climas [Political or digital polarisation? An ecosystem with all climates]. In L. A. Quevedo, & I. Ramirez (Eds.), *Polarizados: ¿Por qué preferimos la grieta? (aunque digamos lo contrario)* [Polarised: Why do we prefer the rift? (Even if we say otherwise)] (pp. 135–154). Capital Intelectual.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.
- Zullo, C. (2023, December 5). *Addressing gender-based violence in the Central Sahel* [Blog post]. InterAction. <https://www.interaction.org/blog/addressing-gender-based-violence-in-the-central-sahel>
- Zygmuntowski, J. J., Zoboli, L., & Nemitz, P. F. (2021). Embedding European values in data governance: A case for public data commons. *Internet Policy Review*, 10(3), 1–29. <https://doi.org/10.14763/2021.3.1572>

Appendix

Introduction

The examination of state-of-the-art research in the Global North and Global Majority World in this book provides a critical evaluation of works focusing on changes in the news media, AI systems, and approaches to governing data, and on the causes and consequences of online mis- and disinformation within information ecosystems. This book is a revised and updated version of *Information and Troubled Democracy: A Global Synthesis of the State of Knowledge on News Media, AI and Data Governance* (Mansell et al., 2025), published by the International Observatory on Information and Democracy, Paris.

This appendix presents the methodology that we followed. We did not undertake a formal “systematic review” following standards such as the Cochrane methodology (Grant & Booth, 2009). The steps we followed in the data collection and analysis phase of our project over two years until October 2024 (updated for this book to May 2025), are described in the following, together with information about the Observatory’s Steering Committee and other contributors to our work.

Themes

The analysis focused on three main themes with a cross-cutting theme.

The *media, politics, and trust* theme focused on understanding the roles of legacy and newer news media in democratic processes and addressed pressures, adaptations, and impacts of changing information ecosystems, including technology, on reporting the news, business models, the political implications of online mis- and disinformation for democracy, and the impacts of mis- and disinformation on trust.

The *artificial intelligence, information ecosystems, and democracy* theme examined the use of AI systems in content moderation and personalisation systems in information ecosystems, their use in legacy and news media production, and the implications for the protection of fundamental rights and for the fabric of democratic societies, including values, structures, and processes, as well as skills requirements, resulting from changes in global communication flows and information quality.

The *data governance and democracy* theme examined the discourses and strategies deployed by Big Tech companies to promote their “AI” solutions in ways that mystify what AI is and can do and how dependence

is created through the exercise of corporate power. AI was approached as an antidemocratic economic and epistemic phenomenon, with particular attention to the implications for democratic processes, practices, and inclusive participation, as well as how data governance in and by actors in the Global North is both exported and resisted.

The cross-cutting *mis- and disinformation* theme provided a focus on the problems for democracy created by the production, circulation, and consumption of information that qualifies under varying definitions of these concepts.

Data collection and analysis

The sourcing of entries in our research corpus was guided by questions set by the Steering Committee (see Table A3), keywords,¹ and bibliographic database searches, including online repositories.² The corpus is based mainly on academic publications, supplemented by reports and other materials from various disciplines and regions. In view of the speed of change and the currency of debates about intersections between corporate and political power, technology, and human beings in today's information ecosystems, "other materials" included preprints, conference papers, and online sources such as blogs, newspapers, and magazines.

Research was included principally from the fields of computer science, engineering, law, media and communications, political economy, political science, psychology, and sociology. The corpus includes works concerned with identifying the impacts of digital technology or mis- and disinformation on society or individuals, and research on the contexts in which information and technology are considered problematic, why, and for whom. Multiple research designs and methods (qualitative, quantitative, experimental, and mixed) are included.

Emphasis was given to sources published from 2020, with exceptions to provide historical context. Titles and abstracts of papers were examined in issues of 52 journals published between 2022 and December 2023, including journals in the social sciences and computer science.³ Google Scholar, Web of Science, and Scopus searches were conducted when sub-topics emerged. Members of Research Assessment Panels (RAPs) and rapporteurs (authors of this book) conducted additional searches. Two global calls for contributions were made to gather sources, especially concerning the Global Majority World, and selected interviews were conducted. Source selection criteria were agreed upon, as shown in Table A1, resulting in a corpus of 3,095 sources (plus 51 additional entries, published between October 2024 and May 2025, added for the book).

Table A1 Corpus source selection criteria

Criterion	Definition
Relevance	Relevance of source to the themes
Publication date	Emphasis on 2022 to October 2024, but earlier sources included as appropriate
Country focus	Global Majority World, Global North, and global (work with global reach)
Document type	Academic peer-reviewed articles, academic books, and book chapters (some books by non-academic authors active in industry), PhD theses/dissertations, research reports, policy reports, conference papers and proceedings, pre-prints, and others (magazine, newspaper articles, and blogs).
Language	English (a small selection of French, German, Italian, Portuguese, and Spanish)
Type of study	Well-established (qualitative or quantitative), social science methodology, and rely on original research or secondary analysis to make an empirical or theoretical contribution. Policy reports or other types of contributions published by reputable organisations that may not adopt a clear methodology or rely on primary data, but identify unique trends.

A set of quality checks was agreed upon, as shown in Table A2 with a view to maximising the inclusion of independent research sources. All selected sources were reviewed throughout the drafting process to examine the public or private (confidential) status of entries and origin (country or public/private organisations). Most sources were in English, but a few entries were in French, German, Italian, Portuguese, or Spanish, and some text was translated using online tools and our own translations. Sources characterised as low or medium quality were not automatically excluded but assessed in the light of differences in the availability of independent funding across countries (i.e., in some countries, research funding is granted mainly by the government or a private foundation). In these cases, the quality of sources was assessed carefully. Sources were assessed for their novelty and robustness in view of their provenance, and they were examined to check whether authors had received financial support, taking note of whether this came from universities, regional or national research funding agencies, companies, civil society, or industry associations when this information was declared.

Table A2 Definitions of source quality

Quality	Definition	Funding
High	Paper published in high-quality, peer-reviewed journal or conference	Independent
	Book/report published by internationally recognised publisher or organisation	
	Book/report written by internationally recognised author(s)	
	Book chapter written by internationally recognised author(s)	
Medium	Paper published in peer-reviewed journal or conference	Independent or not declared
	Book/report published by less well-known publisher or organisation	Independent
	Book chapter written by less well-known author(s)	
	Report	Sponsored (government organisation; intergovernmental organisation; civil society organisation)
Low	Report	Sponsored (private company; private sector association)
	Report	Not declared
	Blog, newspaper or magazine article	–

Weekly meetings were held with the Observatory’s team, scientific director, and rapporteurs (authors of this book), formal reviews of draft text of the earlier report were undertaken by the Steering Committee, and input was received from Research Assessment Panel (RAP) contributors (see Tables A3 and A4 below). The revised and updated text for this book was done independently by the scientific director with editing input from the lead rapporteur after the responsibilities of the Observatory team, Steering Committee, and RAP contributors had ended.

Sources cited

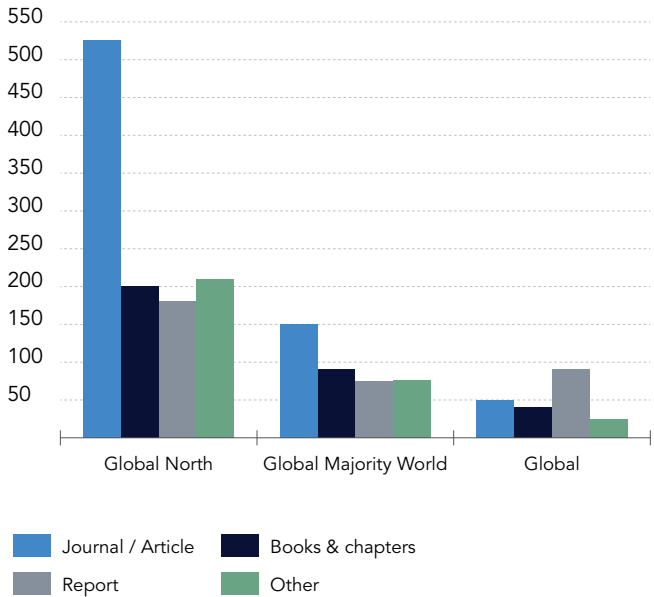
Sources were collated and tagged using Zotero software. Of 1,664 sources cited in the earlier report, 97 per cent were published after 2000; 90 per cent after 2015; and the major proportion published between 2020 and 2024 (43% peer-reviewed articles, 20% books and book chapters, 18% reports, and 18% other materials (conference papers, newspaper and magazine articles, and blog posts).

The regional breakdown of sources distinguished between Global North, Global Majority World, and global. Acknowledging the difficulties of classifying groups of people or countries, the term Global Majority World

was used to refer collectively to the 85 per cent of the world’s population who live in low- and middle-income countries and who are of Indigenous, African, Asian, or Latin American descent (although this sometimes extends to people of dual heritage and to minority ethnic groups who are racialised within countries and not classed as white) (see Anthony et al., 2024; Campbell-Stephens, 2021; Lawrence, 2022; Patrick & Huggins, 2023). Global North was used to refer to those not included in the Global Majority World, generally from the wealthy, industrialised countries. Global refers to works that draw on arguments or evidence from a reasonably large number of countries in the Global North and Global Majority World.

The regional breakdown of the corpus is skewed towards the Global North, making up 65.5 per cent of the 1,664 citations. Despite an effort to address this imbalance, including interviews with experts in the Global Majority World, only 22.5 per cent of citations were either about the Global Majority World or by authors from the Global Majority World. The remaining 12 per cent were global. Figure A1 shows the regional distribution. Many of the reports we examined were produced by independent not-for-profit organisations (we cited 118 of these – 47% Global North, 27% Global Majority World, 26% global). An additional 51 citations were added in this book (55% Global North, 35% Global Majority World, 10% global).

Figure A1 Number of publication types, by regional focus

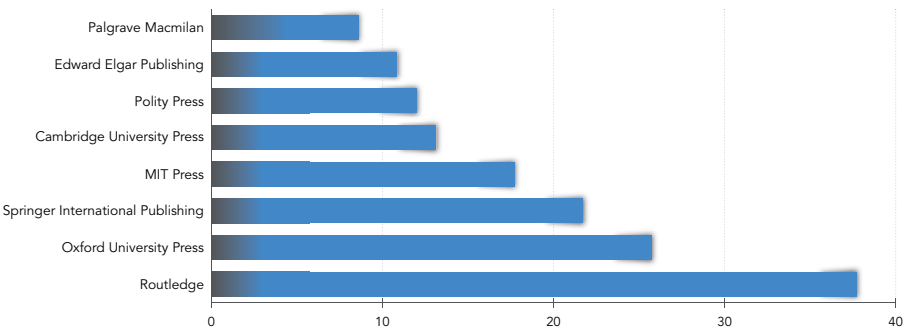


SOURCE: Giovanni Maggi, Observatory on Information and Democracy

NOTE: N = 1,664

Book and book chapter publishers in the corpus are from diverse sources, but mostly based in the United States, the United Kingdom, or Western Europe, and they are often affiliated with academic institutions (e.g., Stanford University, Duke University, University of Amsterdam). In total, books and book chapters are cited from 133 publishing companies. International academic publishing companies are most frequently cited, with the top 20 publishers accounting for 14 per cent of all cited sources ($N = 1,664$) and 37 per cent of all books and book chapters ($N = 332$). The most cited publishers are shown in Figure A2.

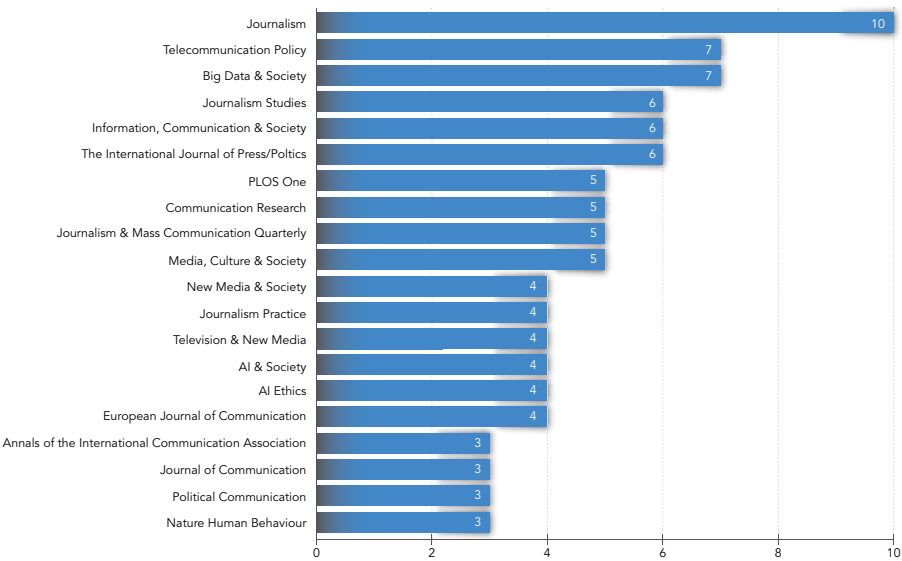
Figure A2 Number of books and book chapters, by most frequently occurring publishing company



SOURCE: Giovanni Maggi, Observatory on Information and Democracy

A total of 712 articles were cited from 132 journals in the earlier report. The most cited were in the media, journalism, and communications fields, as shown in Figure A3, which accounted for 98 citations or 14 per cent of the total journal citations.

Figure A3 Most frequently cited journals



SOURCE: Giovanni Maggi, Observatory on Information and Democracy

Contributors to the analysis

The research assessment was guided by a Prefiguration Group and the Observatory Team, including rapporteurs and a scientific director (this book’s authors), overseen by a Steering Committee (see Table A3), as well as contributors to three thematic Research Assessment Panels (RAPs) and a group of expert external contributors (see Table A4), in collaboration with several hundred others from around the world who are not named here.

Table A3 Research assessment contributors

Group/Team/ Committee	Name	Position
Prefiguration Group	Virgilio Almeida	Professor Emeritus, Universidade Federal de Minas Gerais (Brazil)
	Jim Balsillie	Founder of the Center for International Governance Innovation (CIGI) (Canada)
	Jean-Marie Guéhenno	Diplomat, former United Nations Deputy General Secretary (France)
	Ángel Gurría	Co-chair , former Secretary-General of the OECD
	Miguel Poiares Maduro	Chair, European Digital Media Observatory, European University Institute (Portugal)
	Maria Ressa	CEO of Rappler, Nobel Peace Prize laureate 2021 (Philippines)
	Burhan Sönmez	President, PEN International (Turkey/UK)
	Shoshana Zuboff	Co-chair , Professor Emeritus at the Harvard Business School
Observatory Team	Michael Bak	Executive Director of the Forum on Information and Democracy (France) to April 2024
	Alice Boeffard-Dosierre	Intern, Observatory on Information and Democracy (France)
	Iris Boyer	Head of the Observatory on Information and Democracy (France) to April 2025
	Flavia Durach	Rapporteur News Media, PhD, Associate Professor, Head of Department, Department of Digital Communication and Emerging Media, National University of Political Studies and Public Administration (Romania)
	Chiara Gottardo	Intern, Observatory on Information and Democracy (France) to March 2024
	Camille Grenier	Executive Director, Forum on Information and Democracy (France) from April 2024
	Emma Gruden	Project Officer, Observatory on Information and Democracy (France)
	Matthias C. Kettemann	Rapporteur AI systems, Professor and Head of Department, Department of Theory and Future of Law, University of Innsbruck; Research Program Leader, Alexander von Humboldt Institute for Internet and Society and Leibniz Institute for Media Research, Hans-Bredow-Institute (Austria/ Germany)
	Théophile Lenoir	Rapporteur News Media, PhD candidate, University of Milan and guest doctoral student at the Sciences Po Medialab (France)
	Giovanni Maggi	Research Intern, Observatory on Information and Democracy (France) to April 2025
	Robin Mansell	Scientific Director first research cycle, and lead author for this book, Professor Emerita, Department of Media and Communications, London School of Economics and Political Science (United Kingdom)
	Rob Procter	Lead Rapporteur, Dept. of Computer Science, Warwick University; Alan Turing Institute for Data Science and AI (United Kingdom)
	Gyan Prakash Tripathi	Rapporteur Data Governance, Lawyer and Policy Analyst, (India), Graduate Scholar, Central European University, Vienna (Austria)
	Emily Tucker	Rapporteur Data Governance, Executive Director, Center on Privacy & Technology at Georgetown Law (USA)
	Katharina Zuegel	Policy Director, Forum on Information and Democracy (France)

Continued

Table A3 Continued

Group/Team/ Committee	Name	Position
Steering Committee (Observatory on Information and Democracy, first research cycle)	Luca Belli	Professor, Digital Governance and Regulation, Fundação Getulio Vargas (FGV) Law School, Rio de Janeiro (Brazil)
	Gustavo Cardoso	Professor, Steering Committee Co-Chair ; Media Sociologist, Iscte-IUL; Director of OberCom – Media Observatory (Portugal)
	Niva Elkin-Koren	Professor, Tel-Aviv University Faculty of Law; Faculty Director of the Chief Justice Meir Shamgar Center for Digital Law and Innovation (Israel)
	Helani Galpaya	CEO, LIRNEasia (Sri Lanka)
	Timothy Garton Ash	Professor, European Studies and Isaiah Berlin Professorial Fellow, St Antony's College, University of Oxford; Senior Fellow The Hoover Institution, Stanford University (UK and USA)
	Frances Haugen	Co-Founder Beyond the Screen; algorithmic product expert; advocate for accountability and transparency in social media (USA)
	Jeanette Hofmann	Professor, Freie Universität Berlin; Research and co-founding director Alexander von Humboldt Institute for Internet and Society (HIIG) (Germany)
	Jhalak Kakkar	Executive Director, Centre for Communication Governance, National Law University Delhi; Visiting Professor, National Law University Delhi (India)
	Ansgar Koene	Global AI Ethics and Regulatory Leader, EY (Belgium)
	Admire Mare	Associate Professor, Head Department of Communication and Media, University of Johannesburg (South Africa)
	Nnenna Nwakanma	Digital policy, advocacy and cooperation strategist (Côte d'Ivoire)
	Pier Luigi Parcu	Director, Centre for Media Pluralism and Media Freedom, European University Institute; Director of the Centre for a Digital Society, European University Institute (Italy)
	Courtney C. Radsch	PhD, Steering Committee Co-Chair ; Director, Center for Journalism and Liberty, Open Markets Institute; Fellow UCLA Institute for Technology, Law and Policy; Fellow Center for Democracy and Technology (USA).
	Jeremy Rifkin	Economic and social theorist, bestselling author of 23 books; and President of the TIR Consulting Group LLC (USA)
	Ghassan Salamé	Former Minister of Culture of Lebanon; Emeritus Professor of International Relations, Sciences Po Paris (France and Lebanon)
	Sonja Solomun	Deputy Director, Centre for Media, Technology, and Democracy, McGill University (Canada)
	Nicol Turner Lee	Senior Fellow in Governance Studies and Director Center for Technology Innovation, Brookings Institution, Washington, DC (USA)
	Stefaan Verhulst	Research Professor, New York University (USA); Co-Founder Governance Lab; Co-Founder of Data Tank (Belgium)
	Natalia Andrea Zuazo	Author Guerras de internet and Los dueños de internet; Director of SALTO; UNESCO consultant (Argentina)

SOURCE: Observatory on Information and Democracy

Table A4 Research Assessment Panel (RAP) contributors

AI RAP	Yik Chan Chin (China/UK), Giovanni De Gregorio (Portugal), Paula Gori (Italy), Argyro Karanasiou (UK), Shyam Krishna (UK), Ioannis Kompatsiaris (Greece), Galen Lamphere-Englund (UK), Monica Lopez (USA), Devika Mehta (India), Vincent Obia (UK), Immaculate Odwera (USA), Umut Pajaro Velasquez (Colombia), Anass Sedrati (Sweden), Theodora Skeadas (USA), Sharon Strover (USA), Scott Timcke (South Africa), Evelynne Tauchnitz (Switzerland), Shenja van der Graaf (the Netherlands)
Data Governance RAP	Joren Baillière (Belgium), Poncelet Ileleji (The Gambia), Monica Lopez (USA), Jenna Manhou Fung (Hong Kong SAR), Siddhartha Menon (USA), Umut Pajaro Velasquez (Colombia), Sharon Strover (USA), Michaela Sullivan-Paul (the Netherlands), Sahaj Vaidya (USA)
Media RAP	Anurag (Canada), Jens Barland (Norway), Jose Luis Benitez (El Salvador), Mădălina Boțan (Romania), Paula Gori (Italy), Minna Horowitz (Finland), Ioannis Kompatsiaris (Greece), Jun Liu (Denmark), Beatriz Lopes Buarque (UK), Reyhana Masters (USA), Stephen D. McDowell (USA), Sacha Meuter (UK), Katie Pentney (UK), João Alexandre Peschanski (Brazil), Pauline Renaud (UK), Rose Marie Santini (Brazil), Theresa Josephine Seipp (the Netherlands), Sahaj Vaidya (USA)
External Contributors	Lauro Accioly Filho (Brazil), Theodora Dame Adjin-Tettey (South Africa/Ghana), Susan Ariel Aaronson (USA), Muhammed Alakitan (UK), Gene Allen (Canada), Tsiry Rambel Andrianisa (Madagascar), Olivier Arifon (Belgium), Homa Azodi (Canada), Brian Ball (UK), Habmo Birwe (Senegal), Anita Breuer (Germany), Emma Briant (Australia), Bastien Carniel (France), Kuan-Wei Chen (Taiwan), Lyantoniette Chua (Philippines), Joel Christoph (Italy), Julian D. Cortes (Colombia), Tim Davies (USA), Laurence Dierickx (Norway), Astou Diouf (Senegal), Maira Elahi (Canada), Oscar Espiritusanto (Spain), Jean-Louis Kedieng Ebongue Fendji (Cameroon), Adalberto Fernandes (Portugal), Agustin Ferrari Braun (the Netherlands), Giovana Fleck (the Netherlands), Terry Flew (Australia), Elaine Ford (Peru), Margaret Gallagher (Ireland), Diego Garcia Ramirez (Colombia), Anna Gibson (USA), Rachel Griffin (France), Armando Guio (Colombia), Marci Harris (USA), Sarah Hayes (UK), Petros Iosifidis (UK), Dimitrios Kalogeropoulos (UK), Irini Katsirea (UK), Anselm Küsters (Germany), Aurora Labio-Bernal (Spain), Horacio Larreguy (Mexico), Michel Leroy (France), DeVan Hankerson Madrigal (USA), Sabeehah Mahomed (UK), Fernando Trincado Moraes (Brazil), Alessandro Paciaroni (the Netherlands), Charis Papaevangelou (The Netherlands), Stella Meyer (Belgium), Henna Paakki (Finland), Ryan Payne (Australia), Bethia Pearson (UK), Phil Ramsey (Northern Ireland), Pradipa P. Rasidi (Indonesia), Anna Romandash (Ukraine), Felix M. Simon (UK), Anastasia Stasenko (France), Edward Steinmueller (UK), Zea Szebeni (Finland), Melissa Terras (Scotland), Jorge Vázquez-Herrero (Spain), Karen Vergara Sánchez (Chile), Andrew White (UK), Dolly Wong (Indonesia), Gaston Wright (Argentina), Paul-Alain Zibi Fama (Cameroon)

SOURCE: Observatory on Information and Democracy

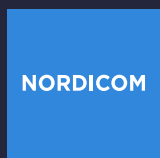
Notes

- ¹ Keywords search in varying combinations: AI, content generation, content governance, content moderation, mediated misinformation, AI literacy, algorithmic transparency, fairness, information behavior, algorithmic bias, democratic institution, machine learning, news personalization, recommender systems, public trust, public sphere, electoral process, algorithmic accountability, AI ethics, deepfake, news media, traditional media, legacy media, digital media, alternative media, public service media, commercial media, digital platform, social media platform, community media, media ecosystem, trust in news, media consumption habits, news avoidance, media diet, journalists, local, regional, national news, monopoly, competition, data governance, epistemic, human rights. These search terms were combined with information ecosystem, information integrity, misinformation, disinformation, malinformation, hate speech.
- ² Online repositories included: AAAI/ACM Conference on AI, Ethics, and Society, AAAI Conference on Artificial Intelligence, ACM Conference on Fairness, Accountability, and Transparency, ACM Conference on Human Factors in Computing Systems, ACM SIGIR International Conference on the Theory of Information Retrieval, Annual Meeting of the Association for Computational Linguistics, European Data Journalism Conference, European Conference on Data and Computational Journalism, IEEE Conference on Artificial Intelligence, International Conference on AI-generated Content, ACM International Conference on Multimedia, International Conference on System Sciences, ArXiv and SSRN Scholarly Papers.
- ³ Keywords for journal search in varying combinations, and mostly excluding Covid-19 papers: information ecosystem, misinformation and disinformation, information integrity, malinformation, hate speech, democracy, algorithm, regulation, and governance.
Journals: *ACM Computing Surveys*, *The African Journal of Information and Communication*, *Artificial Intelligence*, *Artificial Intelligence and Law*, *Artificial Intelligence Review*, *Asian Journal of Communication*, *Asian Journal of Information Technology*, *Berkeley Technology Law Journal*, *Big Data & Society*, *Big Data Mining and Analytics*, *Chinese Journal of Communication*, *Computer Law & Security Review*, *Digital Communication and Networks*, *Digital Journalism*, *Electronic Journal of Information Systems in Developing Countries*, *Energy and AI*, *European Journal of Communication*, *Foundations and Trends® in Machine Learning*, *Frontiers in Big Data*, *Harvard Journal of Law & Technology*, *Human-Computer Interaction*, *Information, Communication & Society*,

Information Development, Information Systems Journal, Innovation and Development, International Data Privacy Law, International Journal of Communication, International Journal of Information Communication Technologies and Human Development, International Journal of Information Management, The International Journal of Press/Politics, International Review of Law, Computers & Technology, Internet Policy Review, Journal of Computer Mediated Communication, Journal of Digital Technologies and Law, Journal of Interactive Advertising, Journal of Law & Innovation, Journal of the ACM, Journalism & Mass Communication Quarterly, Journalism Studies, Knowledge-Based Systems, Law, Innovation and Technology, Media, Culture & Society, New Media & Society, Political Communication, Public Opinion Quarterly, Social Media + Society, Social Science Computer Review, South African Journal of Information Management, Stanford Technology Law Review, Telematics and Informatics, The Information Society.

A critical assessment of why today's information ecosystems are troubling for democracy, this book, based on research in the Global North and Global Majority World, examines changes in news media, artificial intelligence, and data governance. It focuses on measures to improve the quality of public discourse, protect media freedoms, and ensure the resilience of public institutions. The analysis highlights policies aimed at enhancing the accountability and resilience of information ecosystems in the face of the increasing weaponisation of information.

Of interest to academics, policymakers, business, and civil society, the book showcases worldwide efforts to tackle Big Tech data extractive and monetisation practices, address news media instability, and ensure artificial intelligence applications are compatible with democratic values. This book is essential reading for everyone seeking to reimagine what justice in a datafied world could be.



OBSERVATORY ON
INFORMATION AND
DEMOCRACY



Nordicom is funded with
support from the
Nordic Council of Ministers



This book was produced with support from the Observatory on Information and Democracy, and published in collaboration with Nordicom at the University of Gothenburg.