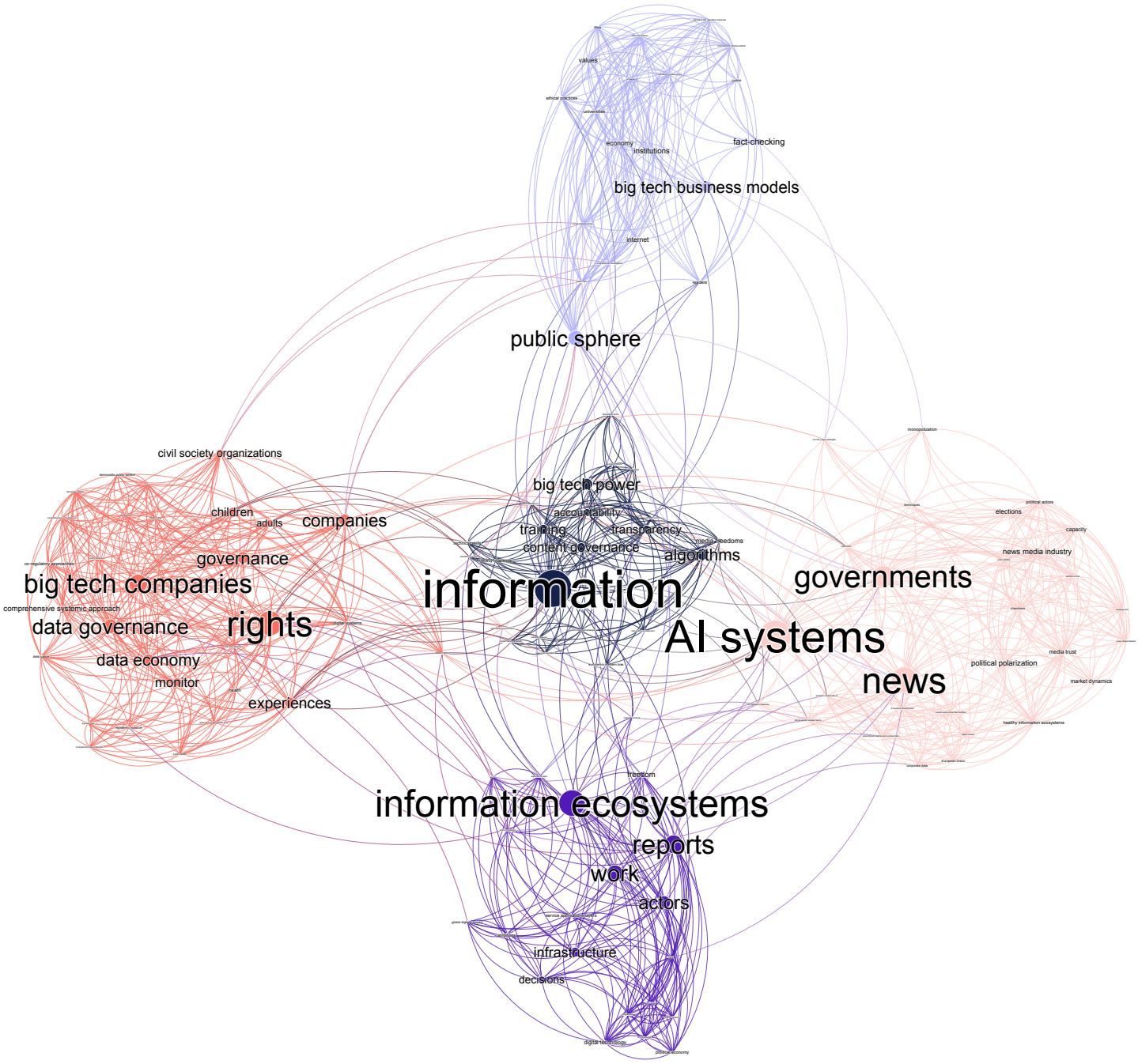
OBSERVATORY ON INFORMATION AND DEMOCRACY

CONCLUSION: INFORMATION ECOSYSTEMS AND TROUBLED DEMOCRACY

An initiative by: FORUM ON INFORMATION AND DEMOCRACY

How to cite this document: Mansell, R., Durach, F., Kettemann, M., Lenoir, T., Procter, R., Tripathi, G., and Tucker, E. (2025) 'Chapter 9: Conclusion: Information Ecosystems and Troubled Democracy' in Information Ecosystems and Troubled Democracy: A Global Synthesis of the State of Knowledge on New Media, AI and Data Governance. International Observatory on Information and Democracy. Paris.



This map represents a statistical summary of the thematic content of this chapter. The network graph represents relations between the words in the chapter, placing them closer to each other the more they are related. The bigger the node, the more present the word is, signalling its role in defining what the report is about. The colors represent words that are closely related to each other and can be interpreted as a topic.

The map is generated by the OID on the basis of the chapter's text using GarganText – developed by the CNRS Institute of Complex Systems. Starting from a co-occurrence matrix generated from chapter's text, GarganText forms a network where words are connected if they are likely to occur together. Clustering is conducted based on the Louvain community detection method, and the visualization is generated using the Force Atlas 2 algorithm.

Link to the interactive map here

'Information and freedom are indivisible'
(Kofi Annan, 1997).[1]

1 Introduction

This report is a critical analysis of state-of-the-art research in the Global North and the Global Majority World that informs us about the interdependent relationships between the cultural, social, political, economic and technological components of information ecosystems. The analysis in the preceding chapters has focused on what interdependence means for the integrity of information and for informed democratic participation in the public sphere.[2] The report addresses three thematic areas with a cross-cutting theme of mis- and disinformation: **media, politics and trust; artificial intelligence, information ecosystems and democracy; and data governance and democracy.**

This research assessment was completed in October 2024, just after the United Nations published a *Global Digital Compact*, highlighting the key challenges and declaring its intent to address them:

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.[3]

From the outset, when work started on our research assessment, we recognized that the design and development, as well as the beneficial and harmful uses of digital technologies, are not simply driven by technological change; they are the result of human decisions and human action. These depend on power relationships embedded in institutions and technologies and how these change over time as a result of the actions of governments, companies, civil society organizations and individuals. For this reason. we understand information ecosystems as a 'system of people, practices, values, and technologies in a particular environment', embedding the public sphere within two layers of the ecosystem: a network infrastructure (hardware and software) layer and a service applications layer.[4]

These actors establish the norms and rules that govern how information ecosystems develop. Rules and norms matter because they affect whether internationally agreed human rights are protected, and whether a public sphere for informed democratic debate will thrive. 'Information and freedom' are indivisible, as former Secretary-General of the United Nations, Kofi Annan said, but information that is complicit in eroding individual rights and collective societal interests in fairness and justice can amount to illegal or harmful mis- and disinformation. When mis- and disinformation are pervasive in the public sphere, we treat this as both a symptom of complex changes in society and as an important amplifier of these changes.

he research assessment was based mainly on academic publications and supplemented by reports and other materials from different disciplines and regions (1,664 citations selected from our bibliographic database, with 3,095 entries screened before inclusion). In view of the speed of change and the currency of debates about

[1] Annan (1997, n.p.), then United Nations Secretary-General.

[2] See Section 3, Chapter 1 for definitions of concepts including information ecosystems, information integrity and the public sphere, and see Appendix: Methodology for the rationale for their use.

[3] UN (2024b, para. 30).

[4] Modified from Nardi & O'Day (1999, p. 49). We acknowledge that there are many ways of defining these ecosystems including those that assume a rhizomatic systems dynamic among the components, see Radsch (2023e).

intersections between corporate and political power and technology and human beings in today's information ecosystems, 'other materials' includes 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. It included works concerned with identifying the impacts of digital technology or mis- and disinformation on society or individuals, and research considering the contexts in which information and technology have come to be seen as problematic, why, and for whom. Multiple research design and methods (qualitative, quantitative and experimental, and mixed methods) were included. The emphasis was on recent published works, rather than an attempt to provide an in-depth account of the history of research in the areas covered.[5]

The decisions and processes involved in undertaking this research assessment, including the steps taken to include work in the Global Majority World (22.5% of citations in this report; Global North 65.5%, Global 12%), selection criteria and quality checks are described in Appendix: Methodology. This global research assessment is not intended to prescribe specific actions for policy makers; rather, it showcases what we can learn from landmark research on the often intractable challenges posed by the rapid changes in information and communication spaces.

Sections 2 and 3 of this concluding chapter discuss central themes that emerged across the chapters of this report. Section 2 highlights insights arising from the analysis of research findings, while Section 3 focuses on the key characteristics of state-of-the-art research, again focusing on insights drawn from a cross-cutting review of the preceding chapters. Section 4 provides a chapter-by-chapter summary, highlighting the core research questions and key findings (readers should go to the individual chapter summaries for a full account of findings and priorities for future research). Section 5 sets out the limitations of the report, and Section 6 concludes with actionable insights that point to what can and could be done to address the 'information crisis'.

2 Principal Thematic Insights

Here we highlight selected themes and conceptual arguments that appear across the topics and issues examined in the chapters of this report (see Figure 9.1). Research communities involved in undertaking studies on questions arising in the thematic areas of interest in this report tend, on the one hand, to welcome the rapid pace of technological innovation and deployment, with a view that any harms will be mitigated, and prioritize getting these technologies to market as quickly as possible for the benefit of humanity. On the other hand, we also draw on insights from research communities that signal the need for caution. In this case, while the many benefits of new technologies (including AI systems) may be recognized, how new technologies become embedded in society and in individual lives is a matter of choice – choice that occurs in a world with unequal power relations.

Figure 9.1
Principal thematic insights

Source: Authors of the report.

Research of this kind typically calls either for greater efforts to introduce mandatory governance for big tech companies or observes that benefits are not being equitably distributed. In this case, the research highlights the need for deeper questioning of the logics of corporate business models and the priorities of those who govern in ways that do

[5] Footnote 1 in Chapters 2 to 8 provides citations for readers seeking background information.

not uphold universal human rights commitments. In brief, the principal thematic insights emerging from this report range from welcoming recent developments in information ecosystems, to skepticism, to outrage.

2.1 HUMAN RIGHTS

Discussions of human rights commitments appear frequently in research on the news media, AI systems and data governance (including governing mis- and disinformation). This includes, but is not restricted to, calls for legislation to protect human rights or to legal interpretations of existing law. Our critical analysis indicates broad agreement that states have a duty to protect human rights and fundamental freedoms, which includes a negative obligation not to violate rights and a positive obligation to protect human rights.

We found no disagreement that internationally protected human rights and fundamental freedoms are pertinent to information ecosystems. In most instances, it is recognized that states need to ensure that their obligations to respect, protect and implement rights are responsive to the challenges posed by information ecosystem actors and instruments. However, it is insisted in the literature that it is essential to differentiate between the normative goals and principles articulated at a global level, and how these are translated over time into practice at local, country and regional levels. Even if international human rights commitments bind signatory states, these are interpreted in different ways, and implementation may not be consistent with normative expectations.

The need to protect media freedoms, freedom of expression and to avoid suppression of voices for political reasons is consistently emphasized. While news media freedom has never been absolute and journalism privileges vary, human rights principles should guide normative expectations for the role of the media, even when there are deviations in practice.

There is substantial evidence that the use of AI systems in content governance can lead to human rights violations. It is well documented that automated content governance and algorithmic decisions can impact negatively on democratic decision-making processes when these systems determine the conditions under which content is seen and with whom it is shared. More generally emerging technologies, such as generative AI (GenAI), challenge both individual rights and rights to democratic participation.

In the literature on data extraction, processing and use, there are several recurrent themes. One is that risk mitigation strategies and practices – voluntary or mandated by legislation – are the preferred means to protect individual privacy, and that corporate appropriation of data generated by people's online interaction is key to prosperous data economies. Another is a questioning of the legitimacy of big tech company data extraction practices and the monetization of data for profit, based on evidence that this leads to unacceptable outcomes, including discrimination and inequalities. In this view, human rights protections are insufficiently robust and the commercial datafication model needs to be reimagined and resisted in the collective interest.

> **Across the issues addressed in this report (media, politics and trust; AI and democracy; and data governance), there is a clear need for research on how international human rights law is interpreted and applied at regional and country level, and whether commitments to protect fundamental rights are actually being met.**

2.2 CONTESTING DATA MONETIZATION

Asymmetrical power relationships and their consequences for strategies and practices of data monetization is a consistent theme in research on changes in the news media industry, AI systems development and use, as well as in research on the role of data in economies. Research repeatedly draws attention to why and how these relationships can lead to disadvantage and discrimination, and the need to acknowledge that these problems arise on both the infrastructure and service applications layers of information ecosystems.

For example, on the infrastructure layer, network neutrality policies and 'zero-rating' contracts are impacting who is connected and who can be disconnected during elections or political unrest, who can access various sources of information, and whether the information ecosystem favors informed participation in the public sphere. This report focused, to a limited extent, on the underlying infrastructure, but it is clear that there needs to be research on the fragmentation or 'balkanization' of the internet, measures to strengthen digital sovereignty and the ambitions of big tech firms and infrastructure service providers alongside research on information integrity problems on the applications layer of information ecosystems. Understanding developments on the infrastructure layer should inform assessments of the health of information ecosystems and acknowledge that the implementation of network neutrality policies and restrictive contracts on data and information access have markedly different effects at different national (and local) contexts.

The news media industry is consistently shown to be influenced by data monetization strategies and AI systems and algorithm developments. These create incentives for legacy news media concentration, destabilize news organizations financially, and lead to closures, especially of smaller local news outlets. Evidence confirms that power asymmetries are at the core of struggles between the news media industry and the big tech company platforms. Power asymmetries are similarly an issue when governments, political parties and other actors manipulate information using datafication (personalization) techniques during critical election periods, and mis- and disinformation are weaponized by both domestic and foreign actors.

> **Our analysis highlights the need for an insight into whether technical competencies are in place to enforce measures to combat harms, especially in times of conflict, but also whether such measures are consistent with a diverse public sphere.**

Asymmetrical power is also visible in research on the monopolization strategies of big tech developers of AI systems and their data practices. In the critical literature, these strategies are found to be misaligned with individual and collective interests, and facilitate the production of mis- and disinformation. Efforts to change these strategies require collective civil society mobilization beyond the capacities of either states or individuals. A repeating theme is that policies that favor the data dependence of private and public organizations as well as individuals are preempting the development of information ecosystems. These include meaningful political deliberation on issues such as rights to data ownership, what role data should have in the private and public sectors, and what contexts require the minimization or prohibition of data production. A further prominent theme is the need for systematic research on data activism initiatives that aim to reimagine ways of restructuring data markets to diffuse concentrations of power that jeopardize democracy.

> **Counter-power strategies would clearly benefit from research aimed at exposing how big tech business models make them attractive targets for mis- and disinformation campaigns, and how digital platforms abandon or arbitrarily change content self-regulatory measures, lay off staff, weaken privacy policies or impose limits on fact-checking.**

2.3 EXCLUSION AND INEQUITABLE INCLUSION

There is a common neglect of the distinctive characteristics of information ecosystems at local, national and regional levels, and especially of differences between the Global North and Global Majority World (and within the Global Majority World). When research considers the impacts of mis- and disinformation, too often the implicit assumption is that these findings can be broadly generalized. Even when this is not the case, as in larger-scale comparative studies that capture impacts mainly at the country level, local and rural/urban experiences are left out of the analysis.

When it comes to assessing the characteristics of trust in news media or in AI systems products and the consequences of how they are infiltrating

people's lives, much of the research we assessed does not consider that internet connectivity is absent for many in the Global Majority World, and meaningful connectivity (affordability, skills and outcomes) is unevenly distributed. News media systems are themselves subject to different ownership, and regulatory regimes and content governance measures can often suppress voices that are critical of state authorities. AI-enabled algorithms promote/demote content in different ways depending on country conditions and political influence. AI systems are deployed in ways that impact communities of color, women, religious minorities and LGBTQ+ people in harmful, yet different, ways. These and other conditions mean that there are substantial differences in how people in low- and middle-income countries, as compared to high-income countries, experience information ecosystems.

In some literature exclusion and inequitable inclusion are discussed with warnings that failure to take difference into account is a recipe for replicating and exacerbating inequalities and injustices. These warnings are present in concerns about individuals, communities and countries becoming dependent on digital infrastructures and algorithmic products produced by big tech companies in the Global North. There is also growing discussion about an 'AI divide' (a growing disparity between those who can access and effectively leverage AI systems and those who cannot). A key theme in our analysis is that homogeneous approaches to governing AI systems and tackling mis- and disinformation are misguided, but there was very little evidence of research on AI system investment strategies being developed that aim to foster international solidarity and inclusive participation.

Inequitable inclusion is also visible in debates in the literature about whether AI systems can be free of bias. Some research insists that they can, but the most prevalent view is that no algorithm or training data set can be free of bias. No content moderation or content curation system can be neutral – there is always the potential for these systems to be used to pursue politicized agendas. As a consequence, outputs of large language models (LLMs) cannot be expected to be fully representative or inclusive on equitable terms, and AI algorithms deployed in the media industry will reflect biases as the result of decisions taken about their design and operation. Some form of epistemic injustice – the privileging of particular kinds of information and knowledge – is always going to be present. The question is how best to counter it.

> **Research points to policies for media freedom, responsible development of AI systems and novel approaches to data governance, but there is a clear need for more work to track the ongoing experiences of Global Majority World countries as they seek to fashion their information ecosystems in ways that are both just and responsive to their conditions.**

Addressing injustices is shown in the literature to require critical thinking about how to govern news media, AI systems and data to counter exclusions and inequitable inclusions. On the policy level, including the *Global Digital Compact*, there are ambitions to tackle exclusions from, and inequitable inclusion in, information ecosystems, and to support measures aimed at enhancing information integrity. In the academic literature there is much discussion of the problems, but little evidence of systematic practical steps to bring about a paradigm shift that would ensure the Global Majority World is not a passive recipient of Eurocentric/Western ideas.

> **There is a clear need to reduce barriers to participation by people in the Global Majority World in all facets of decisions about information ecosystems (including how to treat mis- and disinformation). It is all the more critical to remove these obstacles since they affect the development of AI systems standards and practices. This means devising practices and resourcing them to find creative approaches that ensure that elite Global North knowledge is not the unquestioned guide to governing information ecosystems and the public sphere.**

2.4 TRANSPARENCY AND ACCOUNTABILITY

The concepts of transparency and accountability are pervasively discussed in the literature on governing the news media, AI systems and data, and these appear as being relevant to research on media and information literacy (MIL) and AI literacy.

2.4.1 Governance

A consistent theme is that information ecosystems governance, on the one hand, is too permissive and, on the other, not permissive enough. This varies by topic and by context (whether governance involves democratic or autocratic states). Where governance is found to be too permissive, for example in permitting big tech business to foster the amplification of mis- and disinformation, this is because they are found to privilege their economic self-interest without sufficiently strong rights protections. In short, governance is not strong enough to hold big tech companies accountable, and a lack of transparency in corporate data collection is allowed to persist along with targeted advertising or misuse of data for political gain.

When governance arrangements are found not to be permissive enough, this is typically because states are found to be exerting undue pressure, leading to the suppression of voices. There are considerable differences in views in the literature about how accountability of state and corporate actors is best achieved, that is, through discretionary or mandatory measures. These differences depend on which values receive priority.

> **Our assessment indicates that governance initiatives are needed to tackle the monopolistic power of big tech companies when it is found to unfairly reduce competition and, in some jurisdictions, to lead to harms to privacy. Governance also needs to be strengthened around data collection and to reinforce measures to control stakeholders involved in sharing and selling data. These views are common in both the Global North and Global Majority World, but in the latter, there are concerns about the feasibility of holding distant actors to account.**

Another recurrent theme in the literature is that governance models – for example the European Union's Media Freedom Act of 2024, General Data Protection Regulation (GDPR) or its Artificial Intelligence (AI) Act of 2024 – should not be treated as a panacea for all threats and harms linked to mis- and disinformation. Overall, there is little clarity about the most crucial intervention points where governance can have positive impacts. This means that there is no shared understanding of the appropriate balance between the imperatives of economic growth, innovation and human rights protections when it comes to designing governance to combat mis- and disinformation; this is partly because of concerns about the risks of regulatory overreach by governments, particularly by authoritarian governments.

In the case of measures to promote AI systems transparency and ethical practice in the newsroom and other settings, it is often unclear who is held accountable for harmful outcomes. There are numerous calls in the literature for regular AI systems audits, but less often about who might perform these audits and how they might accomplish them. In the case of AI systems and content governance, there is much evidence of calls for a greater focus on explainability and the development of accountability best practice, but research indicates that the public is unsure about who is responsible for protecting their rights.

> **To hold the big tech companies and governments to account, accurate information is essential in the hands of a wide range of stakeholders. Those whose evidence questions current practice should not be criminalized or marginalized for holding opposing views or for exposing how their interests are not protected.**

2.4.2 Literacy Issues

The challenges created by mis- and disinformation for news media, AI systems and data governance direct attention to promoting enhanced media and information literacy (MIL) (sometimes called digital literacy) as well as AI literacy (data literacy, algorithmic literacy) for individuals (designers of

technology systems, children and adult users of online systems). Much of the literature is concerned with curricula, training and funding, but literacy issues also make an appearance beyond studies of this kind.

Research is clear that MIL and AI literacy policies and initiatives should focus on more than technical skills. Research emphasizes the need to hone critical literacy skills, and for attention to how these skills can be taught effectively to children and adults. Some evidence indicates that those with critical literacy training are less susceptible to mis- and disinformation, although most research only examines the short-term impacts of training and finds a lack of resources, particularly in the Global Majority World. Studies of AI literacy training indicate its necessity at all stages of AI systems development and deployment. It is also clear that literacy initiatives cannot be seen as a solution to all information ecosystems problems, including declining trust in information in the public sphere.

> **MIL and AI literacy should never be presented as a stand-alone project aimed at keeping individuals safe from mis- and disinformation – it must be accompanied by state-led (as appropriate, in view of human rights protections) and individual- or community-led responses to the information crisis.**

Less discussed, but making an appearance in the literature, is that literacy training is also important on a societal level and not just an individual level. Educating the public about the complex issues facing information ecosystems is paramount. An informed public is more capable of demanding accountability from big tech companies and states to ensure that changes in information ecosystems respect human rights. They will be better equipped to insist on the transparency (as far as possible) of algorithmic systems, on human oversight of algorithmic decisions about their lives, and generally, to participate in the public sphere in an informed way. Stronger MIL and AI literacy among policy makers is also essential if they are to devise effective accountability frameworks, monitor and enforce them. We found little evidence on the kinds of research evidence relied on by policy makers in deciding how to govern information ecosystems, counter mis- and disinformation and strengthen democracy.

> **There is little systematic evidence of experience over time on literacy initiatives on a global basis, although the evidence we do have suggests that it can make a positive contribution to individuals' efforts to keep themselves safe online, and to make sense of the information they encounter if training is well resourced.**

3 State-of-the-Art Research and Future Directions

This section provides a critical assessment of state-of-the-art research focusing on consistent themes across the chapters in this report. These themes are related to the Eurocentric/Western bias of research, to the conceptual framing of research, to diverse research design and methods, constraints on researchers' access to data and the independence of research activity (see Figure 9.2).

Figure 9.2
State-of-the-art research assessment

Source: Authors of the report.

3.1 A EUROCENTRIC/WESTERN RESEARCH BIAS

A Eurocentric/Western bias towards research in and on the Global North is as disturbing as it is unhelpful when the aim is to understand the interdependence and power asymmetries of the components of information ecosystems in global contexts. The problems created by mis- and disinformation and approaches to mitigating harms are studied disproportionately in the United States and other Western countries. Large-scale surveys include countries in the Global Majority World, but only some of this work is peer-reviewed, and much of it appears in reports (only some of which are peer-reviewed). The main aim of this research assessment was to examine interactions on a 'system' or institutional level, so we did not include 'case studies' or 'use cases'. If we had included in-depth sectoral or local profiles, there would have been case studies to draw on in regard to news media, although this would have been less likely for uses of AI systems and the challenges of data governance as experienced by civil society groups and activists as well as individual online users beyond the Global North.

Research on which companies – small and large – are involved in datafication processes that yield discriminatory outcomes only focuses on a few companies and does not extend to in-depth assessment of experiences around the world. Longitudinal research on what publics believe should be done about illegal and harmful mis- and disinformation is scarce, and coverage of all countries is missing. We have little insight into which civil society actors are represented in deliberations about how to govern information ecosystems at all levels (local, national, regional and global), and the reasons that they become excluded from these deliberations.

> **The Eurocentric/Western bias of research in all the areas examined in this study needs to be addressed if the views of individuals and organizations in the Global Majority World that work on mis- and disinformation issues are to inform policy in the Global Majority World and debates at international level. A high priority is to work towards decolonizing research and the advice provided to governments and other organizations.**

3.2 CONCEPTUAL FRAMINGS

The use of multiple definitions of concepts across research fields is striking. There is some consistency in the naming of objects of interest and in the way concepts are defined in policy documents, for example information ecosystem, information integrity, mis- and disinformation and 'AI'. These definitions are articulated at an abstract level, and our analysis indicates that their meaning differs in various regions/countries. Despite a shift towards the adoption of several metaphors in recent years, studies emphasize different components of information ecosystems, interpret illegal or harmful information in very different ways, and take different positions on what information should be amplified or suppressed, and whether the focus should be on the public sphere.

In the research community there is debate about whether our object of interest – the information ecosystem – is the priority, or whether the focus should be on the public sphere. There are differences on whether 'information integrity' is too open to varying interpretations on what is good or 'polluting' information. Some prefer to refer to the 'public worthiness' of information to stress informed public discourse and issues such as visibility, access, reflexivity, mediation, influence and legitimacy.

There are few signs of efforts in recent research to conceptualize issues of mis- and disinformation and information integrity in a way that acknowledges lessons from history. Earlier propaganda research is rarely mentioned apart from in studies of trust in news media, where there are some exceptions. In some research 'information society' or 'knowledge society' issues are prominent, although it is not always clear how these concepts differ from what is explored in the case of 'information ecosystems'. There is slippage between how the words 'data', 'information' and 'knowledge' are conceptualized, and lessons from research on the complexity of information environments rarely appear to inform

studies of information ecosystems, except in research that draws on complexity theories of adaptive systems.

Research focuses on different components of information ecosystems without being clear about what is included or excluded. For example, numerous definitions of 'news' appear with a recent strong emphasis on online news that often neglects the role of legacy news media; and research on the role of news media in the public sphere is inordinately focused on news content produced by professional journalists. 'AI' is used – misleadingly – across much of the academic literature that focuses on governance issues and it is used as a generic category of digital systems. This is misleading when the task is to respond to specific risks. In contrast, other research is very specific about the object of study, for example LLMs, but takes little account of the social factors that influence system designs and implementations. Conceptualizations of MIL and AI literacy also differ, and there are no standardized definitions across regions.

Fragmentation of disciplines is common to all academic fields. There are persistent calls for holistic approaches to bridge between the humanities, social sciences and sciences, and to capture the whole lifecycle of mis- or disinformation. Common definitions are important for large-scale, comparative studies, but definitional variety is needed to capture different experiences.

> **In addition to efforts to find common conceptual ground and to be clear about how concepts are defined, efforts to understand how mis- and disinformation are entangled with democracy would benefit from joined-up research with the fields of securitization and the socio-economics of online labor markets. These areas that are not covered in depth in this report, and are rarely cross-referenced in the materials cited in this report.**

3.3 RESEARCH DESIGN AND METHODS

In this report we emphasize that a robust view of mis- and disinformation requires analysis of the complexity of information ecosystems, and this applies to all the objects of interest – the news media industry, AI systems and data governance.

Many studies aim to detect a causal links between mis- and disinformation, changes in attitudes and behaviors and political polarization. Much of this work is informed by theories of media effects, and is undertaken in experimental or quasi-experimental settings or based on respondent self-reporting. The research often points to this kind of information as the 'cause' of political polarization. Other studies point to political knowledge, and whether people belong to homogeneous social, political, cultural and economic groups, as the 'causes' of social discord and distrust that give rise to polarized public opinion. Similarly, some research points to AI systems and algorithms as the 'causes' of changes in attitudes and behavior, and calls for risk mitigation measures. Other studies attribute the causes of instability or conflict to power asymmetries that allow commercial datafication systems to flourish. Studies that find that mis- and disinformation is the cause of filter bubbles and echo chambers are not always sensitive to the conditions in democratic and autocratic regimes.

Fewer studies examine reciprocal relationships between components of information ecosystems (news media, AI systems and datafication processes). Much research focuses on information itself and its impacts, neglecting socio-economic, political and cultural conditions that give rise to it. Also often neglected is the fact that mis- and disinformation are produced and circulated *outside social media*.

> ***Research designs aimed at identifying causal effects of mis- and disinformation on individual attitudes and behaviors needs to be complemented by multidimensional research, on both individual and societal harms, and on the factors in society that give rise to this information.***

Much of the research on countering mis- and disinformation is undertaken in experimental or quasi-experimental settings or based on survey respondent reports, and relies on quantitative

evidence and predictive models. It focuses on the effectiveness of technical tools in providing countermeasures more than on complex factors that give rise to this information. For example, content governance methods such as fact-checking are shown to have effects on people's responses to mis- and disinformation, but most studies are one-off, and unable to account for techniques and practices that change over time.

> **Qualitative (or mixed-methods) research drawing on interviews, focus groups, storytelling, etc., and qualitative data analysis techniques (e.g., thematic, discourse, qualitative content, document analysis) are needed to elicit a deeper insight into complex changes in the public sphere and in the components of information ecosystems, on both the service applications and infrastructure layers.**

Qualitative methods can elicit insight into how power disparities – explicit and hidden – influence choices about the design and deployment of digital technologies and the agency of individuals and groups that engage with these technologies and with digital content. For example, qualitative methods can help to reveal why people value online filter bubbles. Typically treated as having a negative impact on democracy, self-imposed filter bubbles are sometimes valued when they provide a safe space to marginalized groups to express opinions and avoid political or social repression. Qualitative research on AI-driven mis- and disinformation campaigns can provide fine-grained insight into how these processes operate, and why mis- and disinformation is driven by government actors or why it is shared by individuals. It can also help reveal why people's trust in news media and their perceptions of the trustworthy new media organizations vary as much as is indicated by large-scale surveys.

> **Longitudinal comparative studies with global coverage are needed to assess changes in news media trust, political polarization and mis- and disinformation, using both quantitative and qualitative methods. This also applies to research on the role of AI systems and datafication in contributing to the instability of democratic institutions. To advance research that supports healthy information ecosystems and democracy, greater emphasis is needed on interdisciplinary approaches that incorporate diverse research methods and focus on technology's affordances, as well as the practices of states, companies and other key actors.**

3.4 RESEARCHER ACCESS TO DATA

Research in all the fields addressed in this report is limited by problems in accessing real data (in contrast to simulated data). This limits research on decisions in the AI systems development chain, on revenue flows in the news industry and on datafication processes. Legal and ethical issues around the collection and analysis of personal and pseudo-anonymized data create barriers to data collection, and access to government and corporate data is limited in many jurisdictions, although steps are being taken to address this, for example in the European Union.[6] A review of voluntary commitments by OpenAI, Google, Anthropic, Inflection, Meta, Midjourney and Cohere suggests slow progress in providing public application programming interfaces (APIs), deep access to data and policies for researcher access. One study concludes that AI model developers retain exclusive control 'over the majority of research access initiatives'.[7]

[6] Forum on Information and Democracy (2024c). See, for example, Article 40, 'Data access and scrutiny', of the Digital Services Act (DSA), which sets out the conditions for 'vetted researcher' access to data of large online platforms or large online search engines, as designated under the DSA, for research that contributes to the 'detection, identification and understanding of system risks in the Union' and to the 'assessment of the adequacy, efficiency and impacts of the risk mitigation measures' (EC, 2022c). These provisions for designated platforms and the AI Act require data access, but details on data quality are controversial. See Saurwein & Spencer-Smith (2020); van Drunen & Noroozian (2024).

[7] Harrington & Vermeulen (2024, p. 35).

There is an urgent need for safe harbors for researchers and clear data disclosure policies established by data access frameworks. A network of multinational and interdisciplinary research centers dedicated to the study of mis- and disinformation, operating in partnership with online platforms, may help to address the data access problem in countries where frameworks are not in place.

3.5 RESEARCH INDEPENDENCE

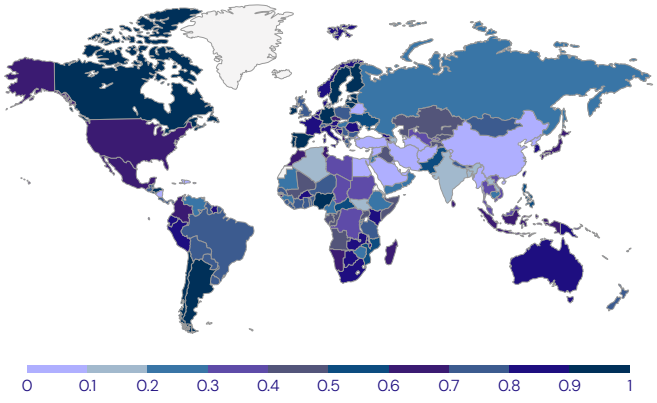
Securing the independence of researchers is vital to the validity and transparency of results. Independence from political interference and from corporate pressures to interpret research findings in ways that favor particular interests, and pressures to prioritize certain research questions, manifests in many ways. Independence is also vital for ensuring that policy makers have access to diverse perspectives and evidence. Our analysis highlights the importance of maintaining the independence of studies of news media trust (and especially studies of the financial sustainability of the press), assessments of the effectiveness of content governance methods, of audits of the performance of AI systems and of data governance frameworks and practices.

In this report we indicate in footnotes the financial support for cited research when it comes from regional/national research funding agencies, companies and a variety of civil society and industry associations when it is declared by the authors of peer-reviewed publications or reports. The intention was not to signal that cited research findings are subject to undue pressure, but that the results should be scrutinized to ensure they are not influenced by interested parties when they are relied upon as evidence.

The independence of research can be challenged when results are deemed to be politically sensitive or to question claims of companies. The politics of conducting research on information ecosystems are revealed when the causes and consequences of online mis- and disinformation are disputed in and outside research institutions.

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'.[8] The overt or subtle silencing of researchers often occurs when the issue is what counts as verifiable knowledge.[9] The reality is that 'academic freedom globally is under threat', with 3.6 billion people living in countries where academic freedom is completely restricted.[10] There is also evidence indicating that increased political polarization correlates positively with levels of academic freedom. Figure 9.3 shows the state of academic freedom in 2023.[11]

Figure 9.3
State of academic freedom, 2023



Source: Kinzelbach et al. (2024, p. 1).
Note: 0-1 scale and color coding indicate low to high freedom

[8] Furstenberg *et al.* (2020).

[9] The European Parliament's Academic Freedom Monitor 2023 notes that, 'From a global perspective, the state of academic freedom in the European Union is relatively high on average compared to other regions and stable over time. Taking the European Union Member States as a reference point, there are nine countries within the European Union with a below-average level of academic freedom. These are Austria, Malta, Romania, Croatia, Bulgaria, the Netherlands, Greece, Poland and Hungary' (EP, 2024, p. 6). Until recently, in Latin America, many universities have retained autonomy, despite authoritarian governments.

[10] Global levels of academic freedom increased to a peak in 2006, but by 2024 declined to their 1973 level. In 2023, only 14.1% of countries were classed as fully free, and 45.5% as completely restricted; see Kinzelbach *et al.* (2024, p. 1), funded in part by the Volkswagen Foundation (VolkswagenStiftung).

[11] Combined with the domination of academic publishing by big publishers, inequalities in knowledge production about mis- and disinformation are likely to escalate, as publishers such as Elsevier, Wiley and Taylor & Francis acquire open access repositories (Posada & Chen, 2018), and as evidence accumulates that scholars are being pressured to retract or revise their conclusions by journal editors (Teixeira da Silva, 2021).

In the *United States*, for example, researchers have been charged with using unethical research methods and results have been censored when election integrity and the role of mis- and disinformation is examined. Claims that research favors certain political parties circulate in the news media and via ostensibly independent organizations, in some cases leading to congressional investigations and undermining public trust in information that circulates in the public sphere.[12] Stanford University's Internet Observatory received pressure for undertaking rapid response tracking of electoral information.[13] A Supreme Court ruling in August 2024 was needed to clarify whether the American government is permitted to communicate with researchers and with social media companies when they undertake research on mis- and disinformation about elections and vaccines.[14] 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.[15]

Controversy about the independent status of researchers and their institutions is not new, and it extends to research on the technical features of information infrastructures when they are deemed to raise national security concerns.[16] Companies such as Google and Meta engage in charitable giving to universities in ways that may be seen as influencing research priorities, or they seed doubt in research findings.[17] Retaliation against employees who become whistleblowers at 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.[18] Research 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'.[19] When academic access to industry AI systems is reduced, this makes it difficult for researchers to interpret industry AI models or to devise public interest alternatives.[20]

> **A recurrent theme is the need to monitor the independence of researchers and their institutions and the impacts of corporate (and in some countries, government) funding.**

4 Chapter Focus and Organization

Here we explain the structure of the chapters in this report. Chapter 1 provided an overview of the core themes and issues addressed, and definitions of the key concepts used. The rest of the report is structured to introduce readers to the research, focusing principally on news media and trust (Chapter 2), the development of AI systems and the implications for human rights (Chapter 3), and the generation and circulation of data within information ecosystems (Chapter 4). The research in these three chapters draws on a variety of theories and empirical evidence on the causes and consequences of changes in information ecosystems, and on the role of mis- and disinformation in changes in the conduct of debate in the public sphere.

Chapters 5 to 8 cut horizontally across the themes to address the public's and policy makers' understanding of issues and controversies and

[12] American Sunlight Project (2024), an independent organization.

[13] In 2024 the US Department of Justice brought a case against Georgia Tech in relation to its cybersecurity lab for refusal to comply with Department of Defense (DoD) security protocols while carrying out DoD-commissioned research; see Abdalla & Abdalla (2021), supported by public money, with a third from industry; DiResta (2024); Menn & Nix (2023); Newton & Schiffer (2024).

[14] Tollefson (2024). The case was initially filed by the then attorneys-general of Missouri and Louisiana, both of whom had challenged whether President Biden had won the 2020 election.

[15] Medronho (2024).

[16] For example, at Georgia Tech security protocols were not implemented for a period of time because they were deemed to compromise the software used in research, leading to charges brought by the Department of Justice; see Anderson (2024); Mueller (2024).

[17] Graham (2024).

[18] Hao (2020); Knight (2024).

[19] Phan *et al.* (2022, p. 130); see also Eastwood (2024).

[20] Ahmed *et al.* (2023, p. 885).

research on literacy training – MIL and AI literacy (Chapter 5), legislative and regulatory measures establishing rules and norms for conduct in information ecosystems (Chapter 6), specific measures (fact-checking, industry self-regulation and co-regulation to mitigate harms of mis- and disinformation (Chapter 7), and individual and collective efforts to imagine and practice data governance consistent with fairness and justice (Chapter 8).

Here are the highlights of Chapters 2 to 8.

Chapter 2: *News Media, Information Integrity and the Public Sphere.* This chapter examined what research tells us about changes in legacy and online news media, and what can be done to promote information integrity and a democratic public sphere. *What are the market structures in the news media industry and the power relations between news media organizations and digital platforms? What is the relationship between news media, a healthy public sphere and democracy? What strategies are available to the journalism profession to work towards building trust in the news media?* The analysis covered research on the structural characteristics of news media markets and platformization, motivations to produce and consume mis- and disinformation and resilience, news media trust and distrust, the trustworthiness of legacy and online news outlets, news consumption and avoidance habits, the weaponization of information and political polarization.

The analysis highlights why market concentration and platform dominance of advertising markets contributes to the financial instability of news media organizations, how these factors affect people's trust (or mistrust) in news media content, and country differences in perceptions of the trustworthiness of news media organizations. Evidence on mis- and disinformation and political polarization is examined, demonstrating inconsistent findings about the causal effects of exposure to, and engagement with, these kinds of content.

The analysis points to the importance of strengthening the bargaining power of news organizations against platforms, differences in findings related to factors influencing news avoidance, and ability to discriminate between accurate and false information. It calls attention to the need to extend research beyond far-right groups to government bodies, ruling political parties and others that manipulate and weaponize information during election periods. Findings are discussed, indicating that self-imposed filter bubbles can help protect marginalized groups by providing a safe space, and based on our review, that partisan online echo chambers are generally found to be smaller than typically assumed in policy debate. It points to the need for research including studies that take account of the role of legacy news media as well as online news media and political actors, longitudinal studies with global coverage to examine changes in media trust and in political polarization, and to independently monitor the news media industry's capacity to sustain trustworthy news.

Chapter 3: *Artificial Intelligence, Information Ecosystems and Democracy.* This chapter examined research on the properties of AI systems (specifically machine learning (ML) algorithms) and their embeddedness in online content governance systems. *How is 'artificial intelligence' (AI) defined, and what are the relationships between AI systems development and internationally protected human rights? What are the interdependencies between AI systems development, the use of automated tools and democratic processes?* The analysis covered research on the relationships between AI systems and human rights, AI systems use and content governance (generation and moderation), and how these developments are related to changes in democracy, societal resilience and cohesion.

This chapter demonstrates how AI systems development and use are co-evolving with the safeguarding of internationally protected human rights and fundamental freedoms. It explains why states need to ensure that their obligations to respect, protect and implement these rights are responsive to specific challenges posed by new actors, instruments and power relations. The analysis highlights that no algorithm or training data set can be free of bias, and that understanding the properties of AI systems is essential if known biases

are to be mitigated. Researchers need to be specific about the algorithms or ML and LLMs that are being examined. There is substantial evidence that the use of AI systems in content governance can lead to rights violations, and modifying content governance practices ignores the multifaceted underlying causes of social discord and distrust that give rise to polarized public opinion.

The analysis concludes that no single content moderation technique will be acceptable to every online participant, which means a strong emphasis is needed on content moderation policies based on multifaceted approaches. AI tools must be used transparently and ethically, and it should not be assumed that AI systems will necessarily enhance newsroom efficiency and productivity. Measuring the scale of mis- and disinformation and its impacts is challenging without access to real data, and there is a need to provide legal frameworks for defining and removing illegal content, assuring accountability and transparency for problematic content, and rules for algorithmic personalization systems. Research needs to include an insight into how human rights law is being interpreted and applied at the country (regional) level, to assess whether commitments to protect fundamental rights are being met, to work on ways to improve data diversity, to research the conduct of (independent) algorithmic audits and, crucially, to address emerging AI divides.

Chapter 4: *Big Tech Power and Governing Uses of Data.* This chapter examined the relationships between the power of big tech companies and approaches to governing data extraction and processes of datafication. *What is the appropriate role of data and digital infrastructures within political communities? How are data aggregation and AI systems changing the way people build, share and receive information and knowledge? How do these big tech strategies and practices interfere with political deliberation, which is essential for the survival of participatory democracy?* The chapter provided an assessment drawing on insights into the political economy of datafication processes, which included research on digital infrastructure contestations, big tech monopolization practices and business models, and the need to work towards democratic forms of data governance.

This chapter reveals injustices associated with the interplay of data extraction and data brokering, and how digital platform business models drive data-intensive economies and a labor market that incentivizes the production of mis- and disinformation. It provides an insight into how powerful (monopolistic) actors within social, economic and political systems determine what data is produced and how it is produced, and the extent to which data is collected in ways that few understand or have control over. It analyses research indicating that much data governance legislation is permissive in fostering the amplification of mis- and disinformation and the entrenchment of global data dependencies. It emphasizes the need for research on how extractive data production has harmful consequences in people's lives, replicating and exacerbating inequalities and injustices. It also addresses data governance frameworks in countries in the Global Majority World that aim to resist the power of big tech companies; on how big tech business models make them attractive targets for mis- and disinformation campaigns; and how online labor markets incentivize the production of mis- and disinformation.

Chapter 5: *Awareness of Mis- and Disinformation and the Literacy Challenge.* This chapter focused on people's knowledge about the presence of mis- and disinformation in information ecosystems and literacy training initiatives aimed at enabling children and adults to identify these types of information and to protect themselves from harmful consequences. *What is known about the scale and severity of mis- and disinformation? How aware are the public and policy makers of the risks and harms of mis- and disinformation? What are the approaches to media and information literacy (MIL), and AI literacy, and what is the evidence on their effectiveness?* It provided an assessment of research in the context of the need to protect the fundamental human rights of both children and adults.

This chapter highlights challenges in measuring the severity of harms of mis- and disinformation to individuals and society in the absence of access to platform data, the tendency to neglect how

conditions in people's offline lives play a role in their experiences of online engagement, and the fact that large-scale studies are limited to a few platforms that are largely centered on the United States.

Research on public awareness of the role of AI systems in generating and circulating mis- and disinformation is shown to reach different conclusions depending on the criteria used and the context, and reveals considerable uncertainty about people's acceptance of interventions by governments or companies to tackle mis- and disinformation. It also neglects analysis of what policy makers understand about the many factors contributing to an 'information crisis'.

It highlights research demonstrating that MIL and AI literacy initiatives need to focus on more than technical skills and should include training in critical literacy; that these initiatives are not a sufficient response to mis-and disinformation; that more research on children's susceptibility to mis- and disinformation is needed to protect the rights of children; and that AI literacy training (and data or algorithmic literacy) are crucial at all stages of AI systems development and deployment. It explains why standardized MIL and AI literacy definitions and cross-country comparative conceptual frameworks and methodologies are needed, and the need for research on how critical literacy skills training can be taught effectively to children and adults.

Chapter 6: *Governing Information Ecosystems: Legislation and Regulation.* This chapter provided an account of selected legislative and regulatory tools available to governments to mitigate the harms of mis- and disinformation, and to govern the way mainly big tech companies operate. *What types of governance approaches are available? What approaches to information ecosystem governance are being promoted at the global level? What are some of the legislative, regulatory and judicial approaches to governing information ecosystems?* This chapter emphasized normative goals and rules embodied in governance approaches, providing an insight into tensions between these goals and their implementation in view of the interests of different actors. The analysis focuses on principles and guidelines reflected in legislation and regulations with respect to network infrastructure, privacy and data protection, digital platforms, AI systems and news media.

This chapter highlights variations in governance measures around the world, especially on the penalization or criminalization of those who produce and circulate mis- and disinformation. It draws attention to research demonstrating why attention to network neutrality policies and 'zero-rating' regulations is crucial, and why human rights principles should guide normative expectations for the role of the news media, even if deviations occur in practice, highlighting that regulation applied to legacy and online news media can result in censorship or leverage over news media organizations. It explains why privacy and data protection legislation is not a panacea for all data economy issues, and why homogeneous approaches to governing AI systems and tackling mis- and disinformation are not likely to be viable. There is a need for research to monitor voluntary and legal governance measures; to track corporate lobbying; to assess whether measures are helping people navigate information ecosystems in ways that enhance resilience to mis- and disinformation; and to assess whether governance is aligned with both individual and collective interests and with experience in the Global Majority World.

Chapter 7: Combating Mis- and Disinformation in Practice. This chapter looked in detail at specific governance measures to combat mis- and disinformation undertaken by civil society organizations and introduced by governments. *What content governance efforts are being made to combat mis- and disinformation? What are the challenges in achieving effective information ecosystems governance? In what ways are human rights protections jeopardized by governance aimed at curtailing online mis- and disinformation? What is known about the public's appetite for interventions to moderate online mis- and disinformation?* The analysis emphasizes the need to differentiate between the stated aims of governance and its consequences when practice falls short of normative expectations. It focuses on fact-checking, industry self-regulation,

co-regulatory approaches and the public's view of how mis- and disinformation issues should be addressed.

This chapter calls attention to research indicating that human rights protections are jeopardized by some measures to combat mis- and disinformation, and the need to achieve greater clarity about intervention points where governance can have positive impacts, illustrating why a single approach is neither feasible nor desirable. There is a need for a shared understanding of the appropriate balance between the imperatives of economic growth, innovation and human rights protections as well as guarding against regulatory overreach, particularly by authoritarian governments. It highlights an overemphasis on technical tools to support mis- and disinformation countermeasures rather than on diverse contexts, emphasizing that practices such as fact-checking are not static processes, so their effectiveness is likely to vary over time. It draws attention to differences between countries in the way they seek to protect press freedom and to counter online mis- and disinformation. Future research is needed to test countermeasures with real-world data beyond Global North countries. The chapter discusses the benefits of using mixed methods to reveal a range of experiences, and the need to monitor digital platform practices that result in the suppression of voices that are critical of state authorities.

Chapter 8: *Towards Data Justice in Information Ecosystems.* This chapter examined research explaining how the monopolistic power of big tech companies creates biases and harmful discrimination and exclusions, infringing on people's human rights in a data economy that thrives on data extraction and monetization. *Why do corporate incentives, strategies and practices involved in designing, developing, selling and controlling data lead to epistemic injustice? What strategies and tactics are individuals and communities developing to resist the extractive features of the data economy?* This chapter emphasizes individual and collective dependencies and inequities resulting from datafication, and how datafication practices might be reimagined to empower individuals and communities in ways that contribute to data justice. It focuses on the consequences of biased AI systems for human rights guarantees and democratic decision making, and individual and group (local, municipal and national) resistance strategies to current practices.

This chapter assesses research demonstrating that commercial datafication supported by AI systems disadvantages and discriminates among people in the data economy by sustaining comprehensive surveillance to enable computerized data production and services. It highlights the epistemic injustices (the privileging of information and knowledge that are neither representative nor inclusive) and the individual and collective dependencies and inequities resulting from datafication, including the consequences of biased data on which AI systems are trained. It reviews research on initiatives taken by individuals and groups to think critically about how to govern massive amounts of digitized data, and highlights strong pressures from civil society to treat data governance as a lever for restructuring data markets, to protect against infringements of human rights and to tackle concentrations of power and wealth that jeopardize democracy. Future research must work on decolonizing knowledge about and experiences of the data economy, monitor discriminatory outcomes of datafication and examine how dependencies on big tech companies are created. It emphasizes the need for greater insight into strategies to advance public interest alternative news media, Indigenous community and municipality initiatives, and develop both community-controlled technologies and decentralized data governance frameworks.

5 Limitations of the Report

This critical analysis of state-of-the-art research on important components of information ecosystems is limited in several ways, which are set out in detail in Chapter 1. Briefly, they include:

- A focus on material inequalities in people's lives only to the extent that broad socio-economic conditions are mentioned, since the analysis is focused on the themes and questions that structured the analysis.

- A principal focus on the upper service applications layer of information ecosystems, although several issues on the infrastructure layer that affect the health of information ecosystems and the quality of debate in the public sphere are discussed.

- An imbalance in Global North and Global Majority World research sources favoring the Global North, notwithstanding our efforts to reach out to be more inclusive.

- The analysis does not aim to cover the extensive research on 'digital divides', although we acknowledge huge variations in the availability of meaningful internet connectivity and access as well the presence of restrictions on access to information.

- This analysis does not cover research on cybersecurity, securitization, geopolitics and 'digital sovereignty' or the economic geography of digital labor markets or the (micro)economic analysis of digital markets.

- The focus tends to be on country-level experience and institutions, with no attempt to include micro-level or sectoral experience, technology 'use cases' or 'case studies'.

- Analysis in this report is inevitably limited by the fact that all research is guided by research questions selected for investigation by research communities, the funding available to do research and the researchers' access to data.

6 A Final Word on what should be done

Our critical analysis of state-of-art-research amply demonstrates a privileging of knowledge about information ecosystems produced in and about the Global North. It also confirms that the affordances of digital systems (including AI systems) are complicit in failures to protect human rights in the Global North and Global Majority World. There is controversy in the research literature about the principal reasons for this – for example, whether the norms and practices of monopolistic companies and states or individual behaviors and attitudes are the predominant explanations for the spread of viral mis- and disinformation. There is an absence of consensus in research evidence about how best to tackle harms associated with mis- and disinformation, and the wider issues around the fragility of democracy. Controversy partly arises from differences in the way problems are identified, conceptualized and studied. Controversy is also attributable to distinctive cultural, social, political and economic conditions in countries around the world.

6.1 ROLE OF CIVIL SOCIETY ORGANIZATIONS

Despite these controversies, this report has identified where future research can help to address them. This report is mainly based on academic research, but it also benefits from research undertaken or commissioned by civil society organizations and other non-university independent non-profits (we cite 118 of these – 47% Global North, 27% Global Majority, 26% Global coverage).

Civil society organizations play a vital role in 'speaking truth to power' and responding to the exploitative data practices of big tech companies and, in some instances, governments. Acknowledging this role involves:

- Recognizing that civil society groups are working with academic researchers to call attention to these practices, and thinking critically about

how to devise just data governance practices and how to build alternative data governance frameworks.

- Encouraging civil society groups and academics who are working on local, community and municipal data governance frameworks and on proposals to introduce decentralized data governance at the national level.

Our report also highlights areas where actions could be taken by governments or the private sector – actions aimed at ensuring that strategies and practices are consistent with international human rights commitments and with strengthening a democratic public sphere. This report was not designed to generate specific recommendations for policy makers or companies, but guidance is provided in this section.

6.2 GUIDANCE FOR POLICY MAKERS

It is important to acknowledge the limits of policy action in the face of corporate power, divided publics, and current political institutional norms and practices, but policy makers can take steps to promote healthier information ecosystems by learning from research evidence. For governments, some actions require new or different governance measures. Others are about how policy makers think about information ecosystem problems, information integrity, the role of new technologies and the problems created by mis- and disinformation.

Tackling Power Asymmetries

- Unhealthy information ecosystems are clearly facilitated by big tech monopolistic business strategies that encourage commercial data monetization. A comprehensive systemic approach is needed if policy makers are to tackle what is widely seen as an 'information crisis' that threatens democratic stability.

- Policy makers should deploy the full range of governance approaches available to them including co-regulatory approaches and competition/anti-trust measures to restrain the big tech industry's use of business models that lead to the amplification of mis- and disinformation and harms to children and adults.

- Evidence indicates that the challenges of governing foreign-owned big tech companies can be addressed by encouraging coalitions of country or regional stakeholders that work to counter the power of these companies.

- Policy must address structural inequalities in digital services markets and political alignments that foster mis- and disinformation which destabilize democracy, especially those that prevent news media independence and stand in the way of treating news media as a 'public good'.

- Steps must also be taken to reimagine and foster alternative datafication models aligned with data justice principles. This means supporting initiatives to build alternative data governance frameworks including local, community and municipal and decentralized national data frameworks and incentivizing the work of civil society organizations that monitor big tech data harmful practices and work to reimagine alternatives.

- It is essential that policy makers preserve and promote the capacities of diverse communities to question dependencies on the products and services provided by big tech companies outside formal policy-making spaces as well as through participation in formal consultative processes.

Independently monitoring human rights infringements

- Investment in monitoring human rights infringements associated with information ecosystems is essential. Evidence indicating that the interests of big tech companies are being favored in policy decisions (even when legislation is in place) due to weak enforcement must lead to steps to put more effective governance in place.

- Policy makers must recognize that measures to combat mis- and disinformation risk suppressing voices that are critical of state

authorities and take steps to ensure that these risks are mitigated.

- Research demonstrates an urgent need for policy measures to secure the safety of journalists and to limit political pressure and other constraints on media freedom especially when these pressures give rise to mis- and disinformation with severe offline impacts.

- Studies highlight the features of datafication strategies that place disproportionate burdens on marginalized populations and these need to be addressed urgently.

Measures to combat mis- and disinformation

- Policy measures are needed to hold big tech companies accountable for the services and AI tools they release to the market. This means monitoring the growing use of personalization systems and AI tools, including GenAI tools.

- Evidence demonstrates that no single content moderation technique will be acceptable to every online participant. This means recognizing that multiple approaches are needed to combat mis- and disinformation, rather than relying disproportionately on AI tools.

- Policy measures are needed to address the financial instability of the news industry in many countries, to promote independent news media and to counter the dependence of news media organizations on digital platforms. This means addressing big tech company resistance to making ad tech revenues transparent, devising ways to ensure independent public service media and smaller local news outlets are financially viable to protect media freedoms and a plural and diverse public sphere.

- Evidence shows that where trust is declining in news media (and public institutions) this cannot be addressed solely by promoting the use of AI systems and other technical measures or by promoting fact-checking. These measures need to be complemented by policies targeting the incentives created by big tech business models.

- It is essential to encourage investigations of the actors and institutions that generate mis- and disinformation and their motivations. These tend to be neglected in policy that favours efforts to mitigate individual harms and a more balanced approach could help to counter the production of mis- and disinformation and its circulation.

Strengthening Transparency and Accountability

- Policy makers must ensure that big tech companies provide fully transparent reports, for example, on content moderation processes (including personalization algorithms), known algorithmic biases, third-party data sharing agreements, and data breaches. This is key to understanding whether these companies are being held to account. Enforcing AI system transparency by ensuring regular independent audits is crucial.

- Research demonstrates that accountability and transparency measures applied to big tech companies can be weakened when barriers exist between state and regulatory institutions charged with implementing them. Policy makers should take steps to improve policy coordination especially for policy aimed at countering mis- and disinformation.

- Evidence indicates that policy coordination is especially important to enforce measures aimed at governing political campaigning and political spending which fosters mis- and disinformation.

Media and Information Literacy (MIL) and AI Literacy

- Media and information literacy (MIL) and AI literacy training for adults and children is a promising means of granting people greater control over their information environment. Policy makers can foster measures to encourage additions to education curricula or encourage private sector and civil society coalitions to provide training and evaluate outcomes over time.

- Recognizing that these training programs are essential to enable children and adults to

interpret and critically the information they are exposed to online is a crucial step.

- It is also essential to evaluate these initiatives over the long term and to adequately resource them.
- MIL or AI literacy training should not be treated as the main solution to unhealthy information ecosystems and declining trust in news media. A systemic approach is needed to address the factors contributing to unhealthy ecosystems.

Influencing Research Priorities

- Frameworks need to be put in place to provide useable data for research purposes. This means implementing frameworks for researcher access to data, ensuring that these are respected, and monitoring concerns of the research community about their adequacy.
- Policy can encourage global cross-disciplinary, collaborative and comparative research through multinational research centers, including the Global Majority World, to examine the incidence and multiple causes of mis- and disinformation. This is likely to require a multinational and interdisciplinary network of research centers operating in partnership with those big tech companies that are willing to acknowledge and examine how data access policies and practices influence what research is undertaken and its results.
- Policy can incentivize multidisciplinary research that joins up work on mis- and disinformation, political processes and market structures with research on cybersecurity and geopolitical tensions. This is essential to capture the interdependency of the components of information ecosystems and their outcomes for individuals and society.
- Policy can encourage research that moves beyond the laboratory to test AI system based methods of detecting and combating mis- and disinformation employing a variety of quantitative and qualitative research methods.
- Policy can help to sustain a public infrastructure for independent research and ensure that research findings are not suppressed for political reasons.
- Policy makers should encourage the inclusion of more diverse types of research in the evidence base that is used to inform policy. This means including research on the socio-economic and political conditions in society that give rise to mis- and disinformation and on the structural and power relations in the big tech industry as well as the impacts on individual attitudes and behaviors.

6.3 GUIDANCE FOR BIG TECH COMPANIES

Big tech companies make public commitments to promote safe and democratic online spaces. These commitments require that companies:

- Change their business strategies and uses of technologies in view of the longer-term negative reputational effects of their data monetization models which harm individuals and groups and are linked to democratic instability.
- Introduce strategies and practices that are fully aligned with international human rights commitments including the rights of the child.
- Provide fully transparent reports voluntarily in countries where legislation is not in place, or in response to legislative requirements, for example, on content moderation processes (including personalization algorithms), third-party data sharing agreements, data breaches, measures to address known algorithmic biases and provide useable data for research purposes.
- Invest in inclusive mechanisms for consulting with individual users and collective organizations about their experiences of mis- and disinformation and their participation in the digitized public sphere and attending to responses that they find acceptable.
- Adequately resource content moderation processes and ensure that the conditions for

workers meet acceptable standards of pay, health insurance and care for mental health.

- Increase transparency by report lobbying expenditure and reporting on which topics – regulatory procedures and court cases – are the target of lobbying activities.

It is crucial to encourage initiatives from all stakeholders – corporate, government, civil society organizations, philanthropic organizations and academics – if the United Nations' *Global Digital Compact* goal of promoting 'diverse and resilient information ecosystems' is to be met. These initiatives will have a greater chance of success if they are based on an understanding of information ecosystems that recognizes their complexity, of how they are developing and experienced differently depending on the context, and on whether government and corporate practices are successful in upholding international human rights commitments.

References

- 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
- 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
- American Sunlight Project. (2024). *The Information Laundering Cycle: How a Coordinated Effort Weaponized the American Political System in Favor of Disinformation*. www.americansunlight.org/ilc-database
- Anderson, N. (2024). After cybersecurity lab wouldn't use AV software, US accuses Georgia Tech of fraud. Ars Technica, 23 August. https://arstechnica.com/security/2024/08/oh-your-cybersecurity-researchers-wont-use-antivirus-tools-heres-a-federal-lawsuit
- Annan, K. (1997). If information and knowledge are central to democracy, they are conditions for development. United Nations Press Release SG/SM/6268, 23 June. https://press.un.org/en/1997/19970623.sgsm6268.html
- DiResta, R. (2024). My encounter with the fantasy-industrial complex. *The Atlantic*, 15 June. www.theatlantic.com/ideas/archive/2024/06/cia-renee-censorship-conspiracy-twitter/678688
- Eastwood, B. (2024). Study: Industry now dominates AI research. MIT Management Sloan School, 18 May. https://mitsloan.mit.edu/ideas-made-to-matter/study-industry-now-dominates-ai-research
- 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
- EP (European Parliament). (2024). *EP Academic Freedom Monitor 2023*. EPRS (European Parliamentary Research Service), STOA (Scientific Foresight Unit), PE 757.798. www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2024)757798
- Forum on Information and Democracy. (2024c). 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
- Furstenberg, S., Prelic, T., & Heathershaw, J. (2020). *The Internationalization of Universities and the Repression of Academic Freedom*. Freedom House US. https://freedomhouse.org/report/special-report/2020/internationalization-universities-and-repression-academic-freedom
- Graham, T. (2024). Is big tech harming society? To find out, we need research – but it's being manipulated by big tech itself. *The Conversation*, 3 October. https://theconversation.com/is-big-tech-harming-society-to-find-out-we-need-research-but-its-being-manipulated-by-big-tech-itself-240110
- Hao, K. (2020). We read the paper that forced Timnit Gebru out of Google. Here's what it says. MIT Technology Review, 4 December. www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru
- 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
- Kinzelbach, K., Lindberg, S. I., & Lott, L. (2024). *Academic Freedom Index Update 2024*. Friedrich-Alexander-Universitat and V-Dem Institute.
- Knight, W. (2024). OpenAI employees warn of a culture of risk and retaliation. *Wired*, 4 June. www.wired.com/story/openai-right-to-warn-open-letter-ai-risk
- Medronho, R. (2024). Far-right seeks to censor science and criminalize researchers studying disinformation in Brazil. Tech Policy Press, 14 June. https://techpolicy.press/farright-seeks-to-censor-science-and-criminalize-researchers-studying-disinformation-in-brazil
- Menn, J., & Nix, N. (2023). Big Tech funds the very people who are supposed to hold it accountable. *The Washington Post*, 6 December. www.washingtonpost.com/technology/2023/12/06/academic-research-meta-google-university-influence
- Mueller, M. (2024). The Justice Department sues Georgia Tech: A teachable moment in cybersecurity management. Internet Governance Project, 27 August. www.internetgovernance.org/2024/08/27/the-justice-department-sues-georgia-tech-a-teachable-moment-in-cybersecurity-management
- Nardi, B. A., & O'Day, V. (1999). *Information Ecologies: Using Technology with Heart*. MIT Press.
- Newton, C., & Schiffer, Z. (2024). The Stanford Internet Observatory is being dismantled. Platformer, 13 June. www.platformer.news/stanford-internet-observatory-shutdown-stamos-diresta-sio
- 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
- 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

- 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
- 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
- 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(2021), 1-6. https://doi.org/10.1016/j.crbeha.2021.100035
- Tollefson, J. (2024). 'Vindicated': Embattled misinformation researchers celebrate key US Supreme Court decision. *Nature*, 26 June. www.nature.com/articles/d41586-024-01766-2
- UN (United Nations). (2024b). *Pact for the Future, Global Digital Compact, and Declaration on Future Generations*. www.un.org/sites/un2.un.org/files/sotf-pact_for_the_future_adopted.pdf
- 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(2024), 1-15. https://doi.org/10.1016/j.clsr.2024.105946