GLOBAL CONFERENCE FOR

MEDIA FREEDOM

NOVEMBER 16, 2020

Freedom of the Media and Artificial Intelligence

Julia Haas, Office of the OSCE Representative on Freedom of the Media

FREEDOM OF THE MEDIA AND ARTIFICIAL INTELLIGENCE

Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development or the Government of Canada. The report is in its original language.

Executive summary

This paper addresses how the use of artificial intelligence (AI) affects freedom of expression and media freedom. While AI can improve communication and information access in numerous ways, including through legacy media, this paper focuses on the main concerns when AI is not deployed in a human rights-friendly manner.

AI can be used as a tool to censor the media and unlawfully surveil citizens and independent journalists. Moreover, in today's online environment, a few dominant internet intermediaries act as gatekeepers in the curation, distribution and monetization of information, including news content. These intermediaries increasingly deploy AI to govern private speech and public discourse.

AI tools, which underpin much of today's content dissemination, are often embedded in the business model of targeted advertising. The use of AI to distribute content based on the predicted preferences of individuals is based on extensive data-driven profiling. To maximize revenue, intermediaries may prioritize content that increases user engagement over providing access to diverse information of public interest or to independent quality journalism. This may undermine users' ability to access pluralistic information and bias their thoughts and beliefs.

To police speech, AI is often applied to identify and remove content considered illegal or undesirable, both by states and intermediaries. The vast amount of available content exceeds the ability for human scrutiny. While AI-based filtering of user-generated content may thus be appealing, AI tools are prone to mistakes. In addition to deploying AI themselves, states mandate private actors to monitor and remove content based on vague definitions within strict timeframes. Such outsourcing of human rights protection to revenue-driven private actors may incentivize over-blocking of legitimate speech and raises additional concerns about the rule of law and discrimination.

AI's potential to facilitate surveillance and censorship for both economic and political reasons poses a threat to the right to seek and receive information, as well as to media pluralism. The power and influence of a few intermediaries, as well as the fact that most AI tools operate opaquely with little regulation or oversight, exacerbates this threat.

This paper also addresses how biases both in datasets and of human developers may risk perpetuating existing inequality, how AI affects legacy media and how the COVID-19 pandemic aggravates the above-mentioned concerns. Providing policy recommendations, this paper concludes that states and the private sector need to guarantee that the design and deployment of AI are grounded in human rights, with transparency and accountability being ensured at all stages.

Freedom of the media and artificial intelligence

Emerging technologies provide unprecedented opportunities for exercising free speech and media freedom.[1] Artificial intelligence (AI) plays an important role in transforming how people communicate and how they consume and engage with media content. AI offers appealing solutions to filter and rank the seemingly infinite user-generated content and information online.[2] As many technological advancements, AI has the potential for good, but can also pose a genuine threat to human rights—in particular, free speech and media freedom.

Despite no universally agreed definition, "AI" is regularly used as an umbrella term for automated, data-driven processes.[3] AI tools may be simple, human-designed instructions; others are more sophisticated and include machine learning. As AI is based on designs and data provided by humans, its outputs are inevitably shaped by cultural values and subjective experiences and beliefs, including inherent biases.[4]

Some states deploy AI to unlawfully surveil citizens and control public communication in ways inconsistent with international human rights law. Enabling unparalleled possibilities for surveillance, AI can facilitate censorship and means to suppress dissent and independent journalism, both online and offline. Consequently, some states use AI to coerce the press and, ultimately, to tighten digital authoritarianism.[5]

Moreover, private actors, in particular providers of search engines and social media platforms, apply AI to filter content in order to identify and remove or deprioritize "undesired" content, known as content moderation, and to rank and disseminate tailored information, referred to as content curation.[6] Both applications regulate speech with the intention to facilitate online communication, provide user-friendly services, and, crucially, increase commercial profit.

AI-powered filtering and ranking of content is enabled by the surveillance of user behaviour at scale. To evaluate and predict the "relevance" of content, AI requires extensive, fine-grained data. These data also facilitate advertising, which is the basis of many internet intermediaries' business model. Commodifying personal data for targeted advertising—which equals profit—incentivizes extensive data collection and processing, a phenomenon described as "surveillance capitalism".[7] Offering services "for free," intermediaries profit from profiling and commercializing the public sphere. Being inherently invasive, this also invites potential abuses of power and pervasive state control.[8] While every form of surveillance has a chilling effect on free speech and the media,[9] AI may impose detrimental constraints on investigative journalists and the protection of sources.[10]

Frequently compared to a "black box,"[11] AI is often opaque and its application invisible.[12] This may lead to the mistaken assumption that its output is neutral and an objective representation of reality. Users may not be aware if AI is utilized, how it obtains a search result or how it promotes or removes content. At the same time, it may not be evident when and how AI is deployed to obstruct the media through surveillance or other forms of interference.[13] Opacity and lack of awareness are major flaws of any AI application.[14]

Opaque AI that governs information dissemination according to business interests may have severe implications for public discourse, particularly in light of the market dominance of very few intermediaries. Oligopolies have become private arbitrators of speech, setting the terms and conditions for global online communication and access to information. Individuals that want to participate in the online sphere are presented with little choice other than to accept the rules and surveillance of dominant intermediaries. Further, such private AI systems and extensive digital footprints may also facilitate state surveillance and political censorship of the press.[15]

The advertising-driven business models at the core of today's internet structure have profoundly affected the sustainability of legacy media by structurally shifting power, to the detriment of quality journalism.[16] The use of AI technologies further shifts this imbalance—with

a particular impact in countries with low internet penetration or no strong public service media.[17]

Any intentional use of AI to interfere with independent reporting—be it through targeted censoring, pervasive surveillance of investigative journalists or using AI-driven bots to attack and silence individual journalists—is a serious threat to media freedom.[18] Even without bad faith, however, the overall use of AI to monitor speech to restrict certain content or disseminate information entails profound risks. While many of the core questions around content removals and curation are not unique to AI, using AI to shape and moderate information at scale exacerbates many existing challenges and gives rise to new ones. The following sections explore the deployment of AI in content moderation and curation, including its potential effects on free speech and media freedom.

Content moderation

The prevalence of certain content, such as violent extremism, hatred or deceptive messages, impairs the quality of public discourse.[19] AI is used to evaluate content in order to flag, demonetize, deprioritize or remove certain content, or ban specific accounts.[20] It is regularly deployed as pre-moderation in the form of upload filters and to analyze content once it is online or after users have reported it.[21] AI will then either take independent action or final assessments remain subject to human review.[22]

AI is still limited in its capability to analyze content. Speech evaluation is highly context-dependent, requiring an understanding of cultural, linguistic and political nuances.[23] Consequently, AI is frequently inaccurate.[24] False positives lead to unjustified limitations on speech, and false negatives may cause a chilling effect, leading to self-censorship and silencing marginalized voices.[25]

Intermediaries' use of AI to proactively moderate content is a form of self-governance, with AI-driven decisions typically based on terms of services or community guidelines.[26] States increasingly request intermediaries to take down specific posts and mandate them to remove certain categories of content, often based on vague definitions, which may lead to blocking news content of public interest.[27] This outsourcing of law enforcement and judicial responsibilities pressures private actors to deploy AI, especially when strict time limits are instructed.[28] While this raises significant concerns regarding the rule of law and due process, it also results in dependence on a few already powerful companies.[29] Altogether, AI seems to accelerate the trend toward general monitoring of communication, which profoundly affects media freedom.[30]

During the COVID-19 pandemic, human moderators under lockdown and an increasing demand to tackle disinformation led to states and intermediaries expanding their reliance on AI. The pandemic illustrated the importance of reliable, pluralistic information, and—as errors increased and remedy responses were delayed—highlighted the need to address AI's own side-effects.[31]

Content curation

With an abundance of online content, user attention is becoming increasingly scarce. Internet intermediaries apply AI to disseminate information that is based on the predicted preferences of individual users.[32] These predictions, however, are driven by intermediaries' intent to monetize data for targeted advertising.[33] Therefore, the AI-fueled curation of newsfeeds and search results seeks to entice users to increase their engagement and time spent on the respective service.[34] Controversial and sensational content can attract more attention, just as misogyny, racism and content instilling fear or hatred can.[35] Hence, AI-powered ranking systems that prioritize "click worthy" rather than newsworthy content may lead to polarization, radicalization and the spread of deceptive or hateful content.[36]

Moreover, increasingly depending on having their content accessed and shared online, legacy media too have to compete for users' attention and may thus be compelled to increasingly focus on "infotainment" rather than on prioritizing public interest, which additionally pressures quality journalism.[37]

The AI-fueled personalization of content, including news, may strengthen users' pre-existing views, creating "echo chambers" and "filter bubbles"[38] and decreasing the likelihood of individuals' exposure to diverse media content.[39] Distorting the perception of reality, this may reinforce power imbalances and amplify "otherness," while seriously threatening media pluralism.[40]

Today's internet structure provides little economic incentive for intermediaries to offer diversity or, indeed, facts. AI that is designed to serve commercial or political interests will unavoidably be biased toward certain types of content in order to nudge and reorient behaviour to optimize profit or persuade, or to intentionally suppress independent journalism.[41] Authoritarians and others can misuse intermediaries' AI systems for nefarious purposes, for example through bots to propagate specific messages or drown out the visibility of journalistic content.[42] AI tools can be used to attack journalists with the aim of silencing them, for example by orchestrating harassment campaigns that simulate a grassroots movement. This method is particularly prevalent in targeting women journalists—and AI-driven distribution systems may even reward such attacks with virality.[43]

While the magnitude of AI's impact on public discourse is still unclear,[44] it is undisputed that AI is regularly deployed to influence people's perceptions and attitudes. Internet intermediaries have become information gatekeepers that use AI to manage media content and information flows, which inevitably shapes users' opinions and behaviour.[45] AI structures can be used to censor the press, by enabling both negative control over information in the form of censorship as well as positive control in the form of propaganda or attacks.[46] Without democratic safeguards, AI-powered monitoring of speech and content dissemination jeopardizes media freedom, access to information and free speech, while at the same time raising concerns about rule of law and systemic discrimination.

Recommendations

People have repeatedly turned to technology to resolve societal challenges. Yet, matters that have long been controversial cannot be resolved solely by outsourcing decision-making processes to AI.[47] Beyond that, technologies can serve as tools for tracking, censorship and repression of the media at an unprecedented scale. While many of the above-mentioned concerns are not unique to AI, its use exacerbates existing threats to free speech and media freedom. To address them effectively, it is crucial to consider the sociotechnical context in which AI is deployed, by whom it is used, and for which purposes. While there can be no one-size-fits-all solution, AI's impact cannot be assessed or addressed in any meaningful way without transparency and accountability.[48]

Having a positive obligation to protect freedom of expression and media freedom, states should promote an environment enabling pluralism.[49] When public authorities deploy AI themselves, they must abide by international human rights standards, ensuring that any restriction of speech or the media is necessary and proportionate.[50] Excessively collecting or merging data in public-private partnerships does not fulfil these criteria. Instead, it often facilitates digital authoritarianism to employ mass surveillance and targeting of individuals and journalists as well as unparalleled censorship.[51] States should not exploit AI to manipulate public opinion, harass journalists or for other repressive ends, but rather to determine acceptable limits on the use of these technologies.

Regulatory measures and AI-related policies should be evidence-based and must not have an adverse impact on media freedom. States should refrain from indiscriminately delegating human rights protection to AI.[52] Furthermore, all endeavours need to be integrated in strong data-protection rules.[53] Consenting to intrusive surveillance practices should not be a pre-condition to participation in online public discourse.

Companies, too, have a responsibility to respect human rights.[54] They should thwart the misuse of their AI systems to suppress dissidents and the press. While many companies commit themselves to "ethics," these are not necessarily in line with human rights.[55] Nevertheless, private initiatives on AI ethics are important, and codes of ethics play a crucial role in corporate social responsibility. Yet, such codes and principles typically lack democratically legitimated safeguards as well as enforcement regimes, and thus they alone cannot provide effective protection.[56]

Transparency is a basic requirement for any public scrutiny.[57] Individuals should know how decisions that affect their lives were produced and which data were processed for what purpose.[58] Regulators and the broader society should know about AI's effects on the media and public discourse. Due to the profound information asymmetry, however, the field remains grossly understudied.[59] Independent research on AI's societal implications should thus be encouraged. To enable scrutiny, AI

needs to be explainable and interpretable.[60] Hence, states should consider making the disclosure of the utilization of AI and its underlying functions mandatory, while being transparent about their own AI deployment. Such requirements could be tiered depending on the specific purpose, actor's role and phase of AI development or application, as well as its risk of violating human rights.[61] Further, clear rules should ensure that information on AI deployment is comparable[62] and that privacy is protected at all phases.[63]

Transparency should go hand-in-hand with increased user agency. Users should have a choice and control over collection, monitoring and analysis of their data for customized content and over intermediaries' interface design.[64] In order to empower users and strengthen citizen's resilience, increased digital literacy is needed.[65]

Transparency is required to know which AI tools are deployed and how automated decisions are made. It is also needed to challenge problematic processes. Those benefiting from AI should be responsible for any adverse consequence of its use. To achieve accountability, strict standards on governance are crucial. Rules should ensure that corporate accountability is tied to companies' profits and that decision makers can be held responsible.[66] States should consider establishing a tiered AI oversight structure,[67] and explore self- and co-regulation models, along with dispute resolution mechanisms, social media councils or e-courts to rapidly determine violations.[68]

To ensure independent scrutiny, national human rights institutions should be empowered to also supervise AI. An important tool is robust human rights impact assessments, which should be conducted periodically throughout the entire AI life cycle and provide publicly available analyses.[69] Moreover, AI tools should be audited regularly and independently,[70] and include careful analysis of whether AI is misused to interfere with the press.

Access to remedies and redress need to be ensured, both for journalists and individual users whose content was restricted by AI, for those who report content, as well as for those harmed by AI-driven interface designs.[71] Remedies must be dealt with in a timely manner and built on sufficient resources.[72] For some automated decision-making processes, human involvement, review and reversibility must be ensured.[73]

Good practices from other fields, including legacy media, can provide lessons to address transparency and accountability.[74]

In addition, the persistent threat of discrimination in both the design and deployment of AI needs to be addressed.[75] Effective responses require a holistic and interdisciplinary approach. Discussions should involve all stakeholders and layers of society, including affected end users,[76] civil society, academia and the media.[77]

Most of the challenges are closely interrelated with the fact that a few dominant companies have significant power and control over the online information ecosystem. A concentration of power, be it by a state or corporation, always entails a risk of far-reaching restrictions of freedoms.[78] States should ensure a competitive AI market to create incentives for alternative business models for intermediary services.[79] Supporting the development of AI tools that are not built on a system of data exploitation and targeted advertising could nurture market pluralism, democratize AI and foster public value-oriented online spaces.[80]

Finally, given the intertwined and transnational nature of AI challenges, it is crucial to join efforts and aim for global solutions. There are various important initiatives, such as those by the Organization for Security and Co-operation in Europe, UNESCO, the Council of Europe or the European Union.[81]

AI is neither a magical bullet for society's challenges nor should it take the blame for all challenges to free speech or media freedom. AI ought not to facilitate digital authoritarianism or high-tech repression of the media. If AI is to enable, rather than undermine, freedom of expression, access to pluralistic information and media freedom, it is imperative for all stakeholders to ensure a human rights-based framework for transparent and accountable AI. As AI increasingly affects every aspect of our communication and media consumption, it is long overdue to embed safeguards in its development and application so that media freedom can thrive.

[1] The right to "… freedom to seek, receive and impart information and ideas … through any media" in enshrined in Article 19 of the International Covenant on Civil and Political Rights (ICCPR) as well as in the UN Universal Declaration on Human Rights and various regional human rights frameworks. The UN Human Rights Council (UNHRC) affirmed that the same rights that people have offline must also be protected online; see resolution on "the promotion protection and enjoyment of human rights on the Internet", A/HRC/RES/20/8, 2012, https://undocs.org/A/HRC/RES/20/8.

[2] It is estimated that in 2020, every minute, users share nearly 150,000 messages and upload 147,000 photos on Facebook, upload 500 hours of videos onto YouTube and post 347,222 stories on Instagram. See Domo, "Data Never Sleeps8.0 ,"https://www.domo.com/learn/data-never-sleeps-8. See also E. Llansó, J. van Hoboken et al., "Artificial Intelligence, Content Moderation, and Freedom of Expression", Transatlantic Working Group (TWG) on Content Moderation Online and Freedom of Expression, 2020, https://www.ivir.nl/publicaties/download/AI-Llanso-Van-Hoboken-Feb-2020.pdf.

[3] See, e.g., OECD, "Recommendation of the Council on Artificial Intelligence", 2020, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449; European Commission's High-Level Working Group on Artificial Intelligence, "A Definition of AI: Main Capabilities and Disciplines", 2019, https://www.aepd.es/sites/default/files/2019-12/ai-definition.pdf.

[4] See, e.g., J. Bryson, A. Theodorou, "How Society Can Maintain Human-Centric Artificial Intelligence,"] 2019, http://www.cs.bath.ac.uk/~jjb/ftp/BrysonTheodorou19.pdf.

[5] See, e.g., United Nations High Commissioner for Human Rights, A/HRC/39/29, 2018, https://undocs.org/A/HRC/39/29; S. Feldstein, "The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression," 2019, https://carnegieendowment.org/files/201901-Feldstein-JournalOfDemocracy.pdf; H. A. Ünver, "Artificial Intelligence, Authoritarianism and the Future of Political Systems," 2018, https://edam.org.tr/wp-content/uploads/2018/07/AKIN-Artificial-Intelligence_Bosch-3.pdf.

[6] See, e.g., B. Bukovska, "Spotlight on Artificial Intelligence and Freedom of Expression #SAIFE," Office of the OSCE Representative on Freedom of the Media (RFoM), 2020, https://www.osce.org/files/f/documents/9/f/456319_0.pdf.

[7] S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York: Public Affairs, 2019.

[8] See, e.g., N. Maréchal et al., "It's the Business Model: How Big Tech's Profit Machine is Distorting the Public Sphere and Threatening Democracy," Ranking Digital Rights, 2020, https://rankingdigitalrights.org/its-the-business-model; J. Gary, A. Soltani, "First Things First: Online Advertising Practices and Their Effects on Platform Speech," Knight First Amendment Institute, 2019; D. Strumpf, W. Fan, "Who Wants to Supply China's Surveillance State? The West," *The Wall Street Journal*, 2017, https://www.wsj.com/articles/who-wants-to-supply-chinas-surveillance-state-the-west-1509540111.

[9] See, e.g., UNHRC, General comment No. 37, 2020, https://www.ohchr.org/EN/HRBodies/CCPR/Pages/GCArticle21.aspx; Privacy International, "Two sides of the same coin," 2018, https://privacyinternational.org/blog/1111/two-sides-same-coin-right-privacy-and-freedom-expression; E. Stoycheff, "Under Surveillance: Examining Facebook's Spiral of Silence Effects in the Wake of the NSA Internet Monitoring," *Journal of Mass Communication Quarterly*, 2016.

[10] J. Posetti, "Protecting Journalism Sources in the Digital Age," UNESCO, 2017, https://unesdoc.unesco.org/ark:/48223/pf0000248054; B. Heller, "Combating Terrorist-Related Content through AI and Information Sharing," TWG on Content Moderation Online, 2019, https://www.ivir.nl/publicaties/download/Hash_sharing_Heller_April_2019.pdf.

[11] F. Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information*, Cambridge, London: Harvard University Press, 2015.

[12] See, e.g., Privacy International, ARTICLE 19, "Privacy and Freedom of Expression in the Age of Artificial Intelligence," 2018, https://www.article19.org/wp-content/uploads/2018/04/Privacy-and-Freedom-of-Expression-In-the-Age-of-Artificial-Intelligence-1.pdf; J. Burrell, "How the Machine 'thinks' Understanding Opacity in Machine Learning Algorithms," *Big Data & Society*, 2016.

[13] See, e.g., E. Thorsen, "Surveillance of Journalists/Encryption Issues," *The International Encyclopaedia of Journalism Studies*, 2019.

[14] See, e.g., K. Yeung, "Responsibility and AI: A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework," Council of Europe, MSI-AUT, DGI(2019)05, 2019, https://rm.coe.int/responsability-and-ai-en/168097d9c5.

[15] Amnesty International, "Surveillance Giants: How the Business Model of Google and Facebook Threatens Human Rights," 2019, https://www.amnesty.org/en/documents/pol30/1404/2019/en.

[16] UN, OSCE, Organization of American States (OAS) and African Commission on Human and Peoples' Rights (ACHRP) mandate holders on freedom of expression and the media, "Joint Declaration on Media Independence and Diversity in the Digital Age," 2018, https://www.osce.org/files/f/documents/1/e/379351.pdf; UNESCO, "Journalism, press freedom and COVID-19," *World Trends in Freedom of Expression and Media Development*, 2020, https://unesdoc.unesco.org/ark:/48223/pf0000373573.

[17] See, e.g., D. Brown, "Big Tech's Heavy Hand Around the Globe: Facebook and Google's dominance of developing-world markets has had catastrophic effects," Human Rights Watch, 2020, https://www.hrw.org/news/2020/09/08/big-techs-heavy-hand-around-globe; A. Hofseth, "E. Bell thinks public service media today has its most important role to play since World War II," NiemanLab, 2018.

[18] See, e.g., "Internet Censorship in China explained," Daxue Consulting, 2020, https://daxueconsulting.com/internet-censorship-in-china.

[19] See, e.g., OSCE project on the safety of female journalists online, https://www.osce.org/fom/sofjo; Report of the UN Special Rapporteur on the promotion and protection on right to freedom of opinion and expression (UN Special Rapporteur), A/HRC/38/35, 2018, https://undocs.org/en/A/HRC/38/35; J. W. Penney, "Internet Surveillance, Regulation, and Chilling Effects Online: A

Comparative Case Study," *Internet Policy Review*, 2017; UN Special Rapporteur, A/74/486, 2019, https://www.ohchr.org/Documents/Issues/Opinion/A_74_486.pdf.

[20] According to Facebook, in the second quarter 2020, 99% of terrorist content and 94.5% of "hate speech" was detected through "proactive detection technology"; see https://transparency.facebook.com/community-standards-enforcement#dangerous-organizations. See also C. Marsden, T. Meyer, "Regulating Disinformation with Artificial Intelligence," European Parliament Research Service, 2019; S. Spandana, "Everything in Moderation: An Analysis of How Internet Platforms Are Using Artificial Intelligence to Moderate User-Generated Content," New America, 2019.

[21] See, e.g., D. Krivokapić, "OSCE RFoM Non-paper on the Impact of Artificial Intelligence on Freedom of Expression," OSCE RFoM, 2020, https://www.osce.org/files/f/documents/b/a/447829.pdf; E. Llansó, "No amount of 'AI' in content moderation will solve filtering's prior-restraint problem," *Big Data & Society*, 2020.

[22] Human review has a high psychological toll on moderators; see, e.g., S. Roberts, "Meet the people who scar themselves to clean up our social media networks," Maclean's, 2018, https://www.macleans.ca/opinion/meet-the-people-who-scar-themselves-to-clean-up-our-social-media-networks; documentary "The Cleaners" by H. Block and M. Riesewieck, 2018.

[23] Council of Europe, "Algorithms and Human Rights: Study on the human rights dimensions of automated data processing techniques and possible regulatory implications," MSI-NET, DGI(2017)12, 2018, https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5. See also, e.g. "Caught in the Net: The Impact of Extremist Speech Regulations on Human Rights Content," EFF, *Syrian Archive*, Witness, 2019, https://syrianarchive.org/en/lost-found/impact-extremist-human-rights#youtube-censorship-of-conflict-documentation-in-syria-yemen-and-ukraine.

[24] For example, Google Jigsaw's Perspective API offers an open-source toolkit to evaluate content's "toxicity". Yet, the team behind it advises against using it for automated moderation due to "too many errors"; see https://conversationai.github.io. See also, e.g., Cambridge Consultants, "Use of AI in Online Content Moderation," Ofcom, 2019; N. Duarte, E. Llansó, "Mixed Messages? The Limits of Automated Social Media Content Analysis," Center for Democracy & Technology (CDT), 2017, https://cdt.org/wp-content/uploads/2017/11/Mixed-Messages-Paper.pdf.

[25] See, e.g., AlgorithmWatch, "Automated moderation tool from Google rates People of Color and gays as 'toxic'," 2020, https://algorithmwatch.org/en/story/automated-moderation-perspective-bias; Fairspeech Project, "The Race Gap in Speech Recognition Technology," 2020, https://fairspeech.stanford.edu.

[26] These are not necessarily in line with human rights standards and can be changed in an opaque way. See, e.g., UN Special Rapporteur, A/HRC/38/35, 2018

[27] See, e.g., UN, OSCE, OAS and ACHRP mandate holders on freedom of expression and the media,"20th Anniversary Joint Declaration: Challenges to Freedom of Expression in the Next Decade," 2019, https://www.osce.org/files/f/documents/9/c/425282.pdf.

[28] Various legislation introduce very short time periods; for example the EU Directive on copyright and related rights in the Digital Single Market, (EU) 2019/790, or the German Network Enforcement Act, 2017.

[29] See, e.g., UN Special Rapporteur, Report on Artificial Intelligence technologies and implications for freedom of expression and the information environment, A/73/348, 2018, https://undocs.org/A/73/348; H. Bloch-Wehba, "Automation in Moderation", *Cornell International Law Journal*, 2020.

[30] States should not impose a general obligation to monitor information that intermediaries transmit; see, e.g., UN Special Rapporteur, A/74/486, 2019, https://undocs.org/en/a/74/486; Council of Europe, Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, 2018, https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=0900001680790e14.

[31] See, e.g., UNESCO, "Journalism, press freedom and COVID-19," 2020; J. York, C. McSherry, "Automated Moderation Must be Temporary, Transparent and Easily Appealable," EFF, 2020, https://www.eff.org/deeplinks/2020/04/automated-moderation-must-be-temporary-transparent-and-easily-appealable; E. Llansó, "COVID-19 Content Moderation Research Letter, CDT, 2020, https://cdt.org/insights/covid-19-content-moderation-research-letter. Relevant information can also be found in internet intermediaries' own reports on COVID-19.

[32] See, e.g., M. DeVito, "From Editors to Algorithms," *Digital Journalism*, 2016; B. Stark, D. Stegmann et al., "Are Algorithms a Threat to Democracy? The rise of Intermediaries: A Challenge for Public Discourse," AlgorithmWatch, 2020, https://algorithmwatch.org/wp-content/uploads/2020/05/Governing-Platforms-communications-study-Stark-May-2020-AlgorithmWatch.pdf; F. Ricci et al., *Recommender Systems Handbook*, Berlin: Springer, 2010.

[33] S. Zuboff refers to a "trade in human futures;" see *The Age of Surveillance Capitalism*, 2019.

[34] See, e.g., C. O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York: Crown Publishers, 2016.

[35] See, e.g., statement by the former Google Chief Technology Officer N. Wong, 2018, https://www.vox.com/2018/9/12/17848384/nicole-wong-cto-lawyer-google-twitter-kara-swisher-decode-podcast-full-transcript, and M. Zuckerberg, "A Blueprint for Content Governance and Enforcement," 2018, https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634.

[36] N. Statt, "Facebook reportedly ignored its own research showing algorithms divided users," *The Verge*, 2020, https://www.theverge.com/2020/5/26/21270659/facebook-division-news-feed-algorithms; J. Cobbe, J. Singh, "Regulating Recommending: Motivations, Considerations, and Principles," *European Journal of Law and Technology*, 2019; M. Ribeiro et al., "Auditing Radicalization Pathways on YouTube," 2019; S. Vosoughi et al., "The spread of true and false news online"; Massachusetts Institute of Technology, 2018, https://science.sciencemag.org/content/359/6380/1146.

[37] N. Helberger et al., "Artificial Intelligence – Intelligent Politics: Challenges and opportunities for media and democracy," Council of Europe, 2020, https://rm.coe.int/cyprus-2020-ai-and-freedom-of-expression/168097fa82; N. Tashidian et al., "Platforms and Publishers: The End of an Era," Tow Center for Digital

Journalism, 2020; E. Bell, T. Owen, "The Platform Press: How Silicon Valley reengineered journalism," *Columbia Journalism Review*, 2017; UNESCO, "Preliminary study on the Ethics of Artificial Intelligence," 2019, https://unesdoc.unesco.org/ark:/48223/pf0000367823; Reporters Without Borders, "Online Harassment of Journalists: Attack of the trolls," 2018, https://rsf.org/sites/default/files/rsf_report_on_online_harassment.pdf.

[38] E. Pariser, "The Filter Bubble," *The Atlantic*, 2010, https://www.theatlantic.com/daily-dish/archive/2010/10/the-filter-bubble/181427; C. R. Sunstein, *Echo Chambers: Bush v. Gore, Impeachment, and Beyond*, Princeton University Press, 2001; B. Bodó et al., "Interested in Diversity: The role of user attitudes, algorithmic feedback loops, and policy in news personalization," *Digital Journalism*, 2018; S. Ignatidou, "AI-driven Personalization in Digital Media," Chatham House, 2019; M. MacCarthy, "The Ethical Character of Algorithms – and What It Means for Fairness, the Character of Decision-Making, and the Future of News," Shorenstein Center on Media, Politics and Public Policy, 2019; K. Yeung, "Five Fears About Mass Predictive Personalisation in an Age of Surveillance Capitalism," *International Data Privacy Law*, 2018.

[39] See, e.g., N. Helberger, "Challenging Diversity – Social Media Platforms and a New Conception of Media Diversity," *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*, New York: Oxford University Press, 2018. See also J. Möller et al., "Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity," *Information, Communication & Society*, 2018; UNESCO, "World Trends in Freedom of Expression and Media Development: Global Report 2017/2018," 2018, https://unesdoc.unesco.org/ark:/48223/pf0000261065_eng.

[40] See, e.g., S. Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York: NYU Press, 2018; UNESCO, Global Report 2017/2018; N. Usher et al., "Twitter Makes It Worse: Political Journalists, Gendered Echo Chambers, and the Amplification of Gender Bias," *The International Journal of Press/Politics*, 2018.

[41] Z. Tufekci speaks about "persuasion architectures" in "We're building a dystopia just to make people click on ads," TEDGlobal, 2017. See also H. Brignull, https://darkpatterns.org; S. Spandana, "Rising Through the Ranks: How Algorithms Rank and Curate Content in Search Results and on News Feeds," New America, 2019, https://www.newamerica.org/oti/reports/rising-through-ranks; Council of Europe, DGI(2017)12, 2018 "Algorithms and Human Rights", https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5]; UN Special Rapporteur, A/73/348, 2018https://undocs.org/en/A/73/348, Report on Artificial Intelligence technologies and implications for freedom of expression and the information environment.

[42] S. Feldstein, "We Need to Get Smart About How Governments Use AI," Carnegie Endowment for International Peace, 2019, https://carnegieendowment.org/2019/01/22/we-need-to-get-smart-about-how-governments-use-ai-pub-78179; C. Cadwalladr, E. Graham-Harrison, "The Cambridge Analytica Files," *The Guardian*, 2018, https://www.theguardian.com/news/series/cambridge-analytica-files; M. Brundage et al, "The Malicious Use of Artificial intelligence: Forecasting, Prevention, and Mitigation," Future of Humanity Institute, OpenAI, Electronic Frontier Foundation et al., 2018.

[43] M. Ressa, "Propaganda war: Weaponizing the internet," Rappler, 2016, https://www.rappler.com/nation/propaganda-war-weaponizing-internet; M. Ferrier, "Attacks and Harassment: The Impact on Female Journalists and Their Reporting," International Women's Media Foundation and TrollBusters, 2018, https://www.iwmf.org/wp-content/uploads/2018/09/Attacks-and-Harassment.pdf; "Online Harassment of Journalists: Attack of the trolls," Reporters Without Borders, 2018, https://rsf.org/sites/default/files/rsf_report_on_online_harassment.pdf.

[44] The research is inconclusive. See, e.g., E. Dubois, G. Blank, "The echo chamber is overstated: the moderating effect of political interest and diverse media," 2018; Stark, Stegmann, "Are Algorithms a Threat to Democracy?" 2020; UNESCO, Global Report 2017/2018.

[45] See, e.g., UN Special Rapporteur, A/73/348, 2018.

[46] N. Wright, ed., "Artificial Intelligence, China, Russia, and the Global Order," Air University Press, 2019, https://www.airuniversity.af.edu/Portals/10/AUPress/Books/B_0161_WRIGHT_ARTIFICIAL_INTELLIGENCE_CHINA_RUSSIA_AND_THE_GLOBAL_ORDER.PDF. S. Feldstein, "The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression," 2019.

[47] This includes questions such as where to draw the line of acceptable speech. See D. Kaye, *Speech Police: The Global Struggle to Govern the Internet*, New York: Columbia Global Reports, 2019. See also, e.g., J. York, C. McSherry, "Content Moderation is Broken. Let Us Count the Ways," EFF, 2019, https://www.eff.org/deeplinks/2019/04/content-moderation-broken-let-us-count-ways.

[48] For civil society initiatives, see, e.g., Santa Clara Principles on Transparency and Accountability in Content Moderation, https://santaclaraprinciples.org; Manila Principles on Intermediary Liability, https://www.manilaprinciples.org.

[49] UN, OSCE, OAS, ACHPR, "20th Anniversary Joint Declaration: Challenges to Freedom of Expression in the Next Decade," 2019.

[50] See Article 19/3 ICCPR, which requires cumulative conditions of legality, legitimacy, and necessity. See also, e.g., UN Special Rapporteur, A/73/348, 2018; Council of Europe, Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems, 2020, https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154.

[51] See, e.g., UNHRC, A/HRC/39/29, 2018; N. Wright, "How Artificial Intelligence Will Reshape the Global Order," *Foreign Affairs*, 2018, https://alinstitute.org/images/Library/ArtificialIntelligenceReshapeGlobalOrder.pdf; UNHRC, General comment No. 37, 2020.

[52] See, e.g., M. Cornils, "Designing platform governance: A normative perspective on needs, strategies, and tools to regulate intermediaries," AlgorithmWatch, 2020, https://algorithmwatch.org/wp-content/uploads/2020/05/Governing-Platforms-legal-study-Cornils-May-2020-AlgorithmWatch.pdf.

[53] See, e.g., Council of Europe, "Guidelines on Artificial Intelligence and Data Protection," Consultative Committee of the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data, 2019, https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8; "Big data, artificial intelligence,

machine learning and data protection," UK Information Commissioner's Office, 2020.

[54] UN OHCHR, "Guiding Principles on Business and Human Rights," 2011, https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_en.pdf; Council of Europe, Commissioner for Human Rights, "Unboxing Artificial Intelligence: 10 steps to protect Human Rights," 2019; UN Special Rapporteur, A/74/486, 2019.

[55] J. Field et al., "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI," Berkman Klein Center for Internet and Society, 2020, http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420; R. Ochigame, "The Invention of 'Ethical AI': How Big Tech Manipulates Academia to Avoid Regulation," *The Intercept*, 2019, https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence.

[56] See, e.g., T. Hagendorff, "The Ethics of AI Ethics: An Evaluation of Guidelines, Minds and Machines," 2020; M. Latonero, "Governing Artificial Intelligence: Upholding Human Rights & Dignity," Data & Society, 2018, https://datasociety.net/wp-content/uploads/2018/10/DataSociety_Governing_Artificial_Intelligence_Upholding_Human_Rights.pdf.

[57] See, e.g., UN Special Rapporteur, A/73/348, 2018; "Spotlight on Artificial Intelligence and Freedom of Expression #SAIFE," OSCE RFoM, 2020; H. Tworek, "How Transparency Reporting Could Incentivize Irresponsible Content Moderation," TWG on Content Moderation Online, 2019, https://www.cigionline.org/articles/how-transparency-reporting-could-incentivize-irresponsible-content-moderation.

[58] Ranking Digital Rights, for example, publishes an annual report on the information disclosed by internet companies; see "2019 RDR Corporate Accountability Index," 2019, https://rankingdigitalrights.org/index2019/assets/static/download/RDRindex2019report.pdf.

[59] B. Bodó et al., "Tackling the algorithmic control crisis—the technical, legal, and ethical challenges of research into algorithmic agents," *Yale Journal of Law and Technology*, 2017; S. Cope, J. York, "Industry Efforts to Censor Pro-Terrorism Online Content Pose Risks to Free Speech," EFF, 2017, https://www.eff.org/deeplinks/2017/07/industry-efforts-censor-pro-terrorism-online-content-pose-risks-free-speech; C. Collett, S. Dillon, "AI and Gender: Four Proposals for Future Research," University of Cambridge, 2019, http://lcfi.ac.uk/resources/ai-and-gender-four-proposals-future-research.

[60] C. Rudin, "Stop explaining black box machine learning models for high stake decisions and use interpretable models instead," *Nature Machine Intelligence*, 2019, https://www.nature.com/articles/s42256-019-0048-x.pdf. The EU General Data Protection Regulation (Regulation (EU) 2016/679) enshrines a right to transparent information; see, e.g. A. Burt, "Is there a 'right to explanation' for machine learning in the GDPR?" *Privacy Tech*, 2017, https://iapp.org/news/a/is-there-a-right-to-explanation-for-machine-learning-in-the-gdpr.

[61] A tiered approach could address opacity due to state or trade secrets and minimize a risk of manipulation of transparent AI. See, e.g., M. MacCarthy, "Transparency Requirements for Digital Social Media Platforms," 2020, https://www.ivir.nl/publicaties/download/Transparency_MacCarthy_Feb_2020.pdf; TWG on Content Moderation Online, "Freedom and Accountability: A Transatlantic Framework for Moderating Speech Online," 2020, https://www.annenbergpublicpolicycenter.org/feature/transatlantic-working-group-freedom-and-accountability; UN, OSCE and OAS mandate holders on freedom of expression and the media, "Joint Declaration on Freedom of Expression and Elections in the Digital Age," 2020, https://www.osce.org/files/f/documents/9/8/451150_0.pdf.

[62] Current transparency reports do not provide meaningful insights and are not comparable to one another. See, e.g., D. Keller, P. Leerssen, "Facts and Where to Find Them: Empirical Research on Internet Platforms and Content Moderation," in: N. Persily, J. Tucker, *Social Media and Democracy: The State of the Field and Prospects for Reform*, Cambridge: University Press, 2019; P. Leerssen, "The Soap Box as a Black Box: Regulating Transparency in Social Media Recommender Systems," Institute for Information Law, University of Amsterdam, 2020.

[63] UNHRC, Resolution "The right to privacy in the digital age," A/HRC/RES/34/7, 2017, https://undocs.org/A/HRC/RES/34/7; UNHRC, A/HRC/39/29, 2018; UN Special Rapporteur, A/HRC/29/32, 2015, https://undocs.org/en/A/HRC/29/32.

[64] J. Harambam et al., "Designing for the Better by Taking Users into Account: A Qualitative Evaluation of User Control Mechanisms in (News) Recommender Systems," 2019, https://www.ivir.nl/publicaties/download/paper_recsys_19.pdf ; UN Special Rapporteur, A/73/348, 2018.

[65] See, e.g., Council of Europe, Recommendation CM/Rec(2019)10 of the Committee of Ministers to member States on developing and promoting digital citizenship education, 2019, https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=090000168098de08.

[66] See, e.g., "Getting to the Source of Infodemics: It's the Business Model," Ranking Digital Rights, 2020; Access Now, "Protecting Free Expression in the Era of Online Content Moderation," 2019, https://www.accessnow.org/cms/assets/uploads/2019/05/AccessNow-Preliminary-Recommendations-On-Content-Moderation-and-Facebooks-Planned-Oversight-Board.pdf.

[67] See, e.g., C. Marsden, T. Meyer, "Regulating Disinformation with Artificial Intelligence," 2019; C. Wendehorst, C. Woopen, "Opinion of the German Data Ethics Commission," 2019, https://www.bmjv.de/DE/Themen/FokusThemen/Datenethikkommission/Datenethikkommission_EN_node.html; M. Wieringa, "What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability," ACM Conference on Fairness, Accountability, and Transparency, 2020.

[68] TWG on Content Moderation Online, "Freedom and Accountability," 2020; H. Tworek et al., "Dispute Resolution and Content Moderation: Fair, Accountable, Independent, Transparent, and Effective," 2020, https://www.ivir.nl/publicaties/download/Dispute_Resolution_Content_Moderation_Final.pdf; ARTICLE 19, "Social Media Councils," https://www.article19.org/social-media-councils. For self-regulatory initiatives, see, e.g., Global Network Initiative, https://globalnetworkinitiative.org; IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, https://ethicsinaction.ieee.org; Partnership for AI, https://www.partnershiponai.org.

[69] The UN Guiding Principles on Business and Human Rights require human rights due diligence, and that actual and potential human rights impacts are identified and addressed. See also D. Reisman et al., "Algorithmic impact assessments: a practical framework for public agency

accountability," *AI Now*, 2018, https://ainowinstitute.org/aiareport2018.pdf; A. Mantelero, "AI and Big Data: A blueprint for a human rights, social and ethical impact assessment," *Computer Law & Security Review*, 2018.

[70] The Women Leading in AI, for example, propose introducing certificates for audited AI systems; see "Women Leading in AI: 10 Principles of Responsible AI," 2019, https://womenleadinginai.org/wp-content/uploads/2019/02/WLiAI-Report-2019.pdf. See also J. Kroll et al., "Accountable Algorithms," *Pennsylvania Law Review*, 2017, https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=9570&context=penn_law_review; C. Sandvig et al., "Auditing Algorithms," 2014, http://social.cs.uiuc.edu/papers/pdfs/ICA2014-Sandvig.pdf; H. Ziady, "Facebook and YouTube accept hate speech audits to keep advertisers happy," *CNN Business*, 2020, https://edition.cnn.com/2020/09/23/tech/facebook-youtube-advertisers/index.html?mc_cid=eb7d5959b9&mc_eid=8db8161786.

[71] See, e.g., Report of the UN Special Rapporteur on freedom of expression, A/HRC/32/38, 2016; UN, OSCE, OAS and ACHRP mandate holders on freedom of expression and the media, "Joint Declaration on Freedom of Expression and 'Fake News,' Disinformation and Propaganda," 2017, https://www.osce.org/files/f/documents/6/8/302796.pdf; J.York, C. McSherry, "Content Moderation is Broken. Let Us Count the Ways", 2019.

[72] For remedies and redress to be effective, resources and leadership commitment is crucial; see, e.g., C. Silverman et al., "'I Have Blood on My Hands': A Whistleblower Says Facebook Ignored Global Political Manipulation," *BuzzFeed News*, 2020, https://www.buzzfeednews.com/article/craigsilverman/facebook-ignore-political-manipulation-whistleblower-memo.

[73] See, e.g. K. Bontcheva et al., "Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression," International Telecommunications Union (ITU) and UNESCO, 2020, https://www.broadbandcommission.org/Documents/working-groups/FoE_Disinfo_Report.pdf; R. Gorwa et al., "Algorithmic content moderation: Technical and political challenges in the automation of platform governance," *Big Data & Society*, 2020; G. Wang, "Humans in the Loop: The Design of Interactive AI Systems," 2019, https://hai.stanford.edu/blog/humans-loop-design-interactive-ai-systems. The UN Special Rapporteur called to ensure a "human in the loop" in AI tools (A/74/486, 2019).

[74] Regarding legacy media, editorial decision-making typically needs to be independent from commercial operations. Other interesting fields could be product liability; see European Commission, "White Paper on Artificial Intelligence," or aerospace regulation; see I. Raji et al., "Closing the AI accountability gap," 2020. Lessons can also be drawn from environmental law (impact assessment), car regulations (safety and speed limits), medicine (certifications), military ("meaningful human control"), telecom regulation, etc.; see J. Ausloos et al., "Operationalizing Research Access in Platform Governance: What to learn from other industries?," AlgorithmWatch, 2020, https://algorithmwatch.org/wp-content/uploads/2020/06/GoverningPlatforms_IViR_study_June2020-AlgorithmWatch-2020-06-24.pdf.

[75] See, e.g., Amnesty International et al., "Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems," 2018, https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf; Lighthouse3 initiative, list of women in AI ethics, https://lighthouse3.com/womeninaiethics; AI Now, "2019 Report," 2019, https://ainowinstitute.org/AI_Now_2019_Report.pdf; UNESCO, "Artificial Intelligence and Gender Equality," 2020, https://en.unesco.org/AI-and-GE-2020; Council of Europe, Recommendation of the Committee of Ministers to member States on media pluralism and transparency of media ownership, CM/Rec(2018)1, 2018, https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680790e13; R. Avila et al., "Artificial Intelligence: open questions about gender inclusion," WWW Foundation, http://webfoundation.org/docs/2018/06/AI-Gender.pdf.

[76] A recent study by BEUC, the European Consumer Organisation, found that users do not trust AI applications with regard to their privacy and are concerned AI is used to manipulate their decisions. Only 20% found existing legislation adequate. See BEUC, "Artificial Intelligence: what consumers say," 2020, https://www.beuc.eu/publications/beuc-x-2020-078_artificial_intelligence_what_consumers_say_report.pdf.

[77] X. Hu et al., "Steering AI and advanced ICTs for knowledge societies: A Rights, Openness, Access, and Multi-stakeholder Perspective" UNESCO, 2019, https://unesdoc.unesco.org/ark:/48223/pf0000372132.locale=en; N. Helberger et al., "Governing online platforms: From contested to cooperative responsibility," *The Information Society*, 2017; Council of Europe, CM/Rec(2018)1, 2018; F. Carugati, "A Council of Citizen Should Regulate Algorithms," WIRED, 2020, https://www.wired.com/story/opinion-a-council-of-citizens-should-regulate-algorithms.

[78] Regarding legacy media, see, e.g., UNHRC, General comment No. 34 on preventing "undue media dominance or concentration of privately controlled media groups in monopolistic situations that may be harmful to a diversity of sources and views." UN, OSCE, OAS, ACHPR, "20th Anniversary Joint Declaration: Challenges to Freedom of Expression in the Next Decade," 2019; Z. Teachout, *Break 'Em Up: Recovering Our Freedom from Big Ag, Big Tech, and Big Money*, New York: St. Martin's Press, 2020.

[79] Ideas for more competitiveness include norms on interoperability, data portability, decentralization, open source requirements, etc. See, e.g., Council of Europe, CM/Rec(2020)1, 2020; J. Furman et al., "Unlocking digital competition: Report of the Digital Competition Expert Panel," 2019, https://www.gov.uk/government/publications/unlocking-digital-competition-report-of-the-digital-competition-expert-panel; U. Dolata, *Internet Companies: Market Concentration, Competition and Power*, Berlin: Springer, 2018.

[80] Public value-oriented online spaces could, for example, provide a public utility service for indispensable communication. See, e.g., E. Zuckerman, "The Case for Digital Public Infrastructure," Knight First Amendment Institute, 2020, https://knightcolumbia.org/content/the-case-for-digital-public-infrastructure; I. Moura, "AI is transforming society: Here's what we can do to make sure it prioritizes human needs," Medium, 2020, https://medium.com/human-machine-collaboration/ai-is-transforming-society-heres-what-we-can-do-to-make-sure-it-prioritizes-human-needs-34794565aad0.

[81] See, in particular, OSCE, Impact of Artificial Intelligence on Freedom of Expression, https://www.osce.org/fom/ai-free-speech; UNESCO, "Elaboration of a Recommendation on

the ethics of artificial intelligence," https://en.unesco.org/artificial-intelligence/ethics, Council of Europe, "Council of Europe and Artificial Intelligence," https://www.coe.int/en/web/artificial-intelligence; European Union, "Artificial Intelligence," https://ec.europa.eu/digital-single-market/en/artificial-intelligence; ITU, "Artificial Intelligence," https://www.itu.int/en/ITU-T/AI/Pages/default.aspx; and OECD, "OECD Principles on AI," https://www.oecd.org/going-digital/ai/principles.