Mollen, Anne, 2025. "Struggling with generative AI: Digital self-determination along infrastructures of automation."
communication +1, vol. 11, issue 2, pp. 1—30.
DOI: https://doi.org/10.7275/cpo.2247

Struggling with generative AI: Digital self-determination along infrastructures of automation

Anne Mollen, University of Münster, GER, anne.mollen@uni-muenster.de

The widespread adoption of generative artificial intelligence (AI) has led to struggles on a global scale by individual and collective actors trying to secure their autonomy with reference to generative AI. There are several examples of how generative AI impacts the ability of individuals and collectives to self-govern and exercise their free will: for instance, training data copyright violations, cultural misrepresentations, precarious working conditions of data workers, and the environmental and social justice implications of generative AI development. Based on discussions from various fields, including machine learning, AI ethics, and critical data studies, this article presents how current struggles in generative AI relate to matters of autonomy, sovereignty, and self-determination. It specifically reflects on autonomy in relation to generative AI training data, accountability, and market concentration as well as social and environmental justice. Given that these struggles over autonomy significantly relate to the materiality of generative AI, the article proposes digital self-determination and an infrastructural perspective as an analytical concept for a multi-actor, process-oriented, and situated contextual analysis of how autonomy implications manifest in generative AI infrastructures.

Introduction: Struggling with generative AI

The launch of ChatGPT in November 2022 not only led to widespread extensive publicity about artificial intelligence (AI) capabilities but has also been accompanied by profound struggles surrounding generative AI as a technology. Discussions of the potency of generative AI systems, their safety, and possible existential risks to humanity have led to discussions about hypothetical AI risks and attempts to align AI systems with human intentions (AI safety/alignment).[1] Another strand of discussion has focused on the already prevalent risks and harms posed by AI systems—specifically large language models (LLMs) that build the basis for most current generative AI systems—such as bias and discrimination, their ecological impact, and privacy or data violations (AI ethics).[2] Some examples illustrating the current struggles that ensued after the widespread application of generative AI systems include competition and market concentration, copyright violations, discriminatory outputs, precarious working conditions along the LLM value chain, and political misinformation.[3]

These struggles exemplify the attempts of individuals, collectives, and societies to secure autonomy—in the sense of the freedom to act independently and not having their actions determined by others—in relation to LLMs. They constitute attempts to secure and uphold certain principles and values in individuals' personal lives and societies in a situation in which fundamental shifts and challenges to established values occur through generative AI.

Many of these struggles that followed the launch of ChatGPT and similar applications already had a trajectory within critical machine learning (ML) research,

---

[1] Yoshua Bengio et al., "Pause Giant AI Experiments: An Open Letter," *Future of Life Institute*, March 22, 2023, https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

[2] Emily M. Bender et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21* (Association for Computing Machinery, 2021), 610-23, https://doi.org/10.1145/3442188.3445922; Solon Barocas et al., "Fairness and Machine Learning: Limitations and Opportunities," fairmlbook.org, 2019; Henrik Skaug Sætra and John Danaher, "Resolving the Battle of Short- vs. Long-Term AI Risks," *AI and Ethics*, September 4, 2023, https://doi.org/10.1007/s43681-023-00336-y.

[3] Irene Solaiman et al., "Evaluating the Social Impact of Generative AI Systems in Systems and Society," preprint, arXiv, June 12, 2023, https://doi.org/10.48550/arXiv.2306.05949; Marie-Therese Png, "At the Tensions of South and North: Critical Roles of Global South Stakeholders in AI Governance," in *ACM Conference on Fairness, Accountability, and Transparency, FAccT '22*, (Association for Computing Machinery, 2022), 1434–45, https://doi.org/10.1145/3531146.3533200; David Gray Widder et al., "Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI," *SSRN Electronic Journal* (2023), https://doi.org/10.2139/ssrn.4543807. Mary L. Gray and Suri Siddharth, *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass* (Houghton Mifflin Harcourt, 2019).

for instance, on algorithmic fairness as well as the monopolistic, hegemonic, and extractivist tendencies in LLM production.[4] The response in such engineering-oriented fields has often been a technical solution; for instance, how to better determine societal risks of generative AI through assessments[5] or transparency[6] and how to develop algorithms more fairly[7] and in a more environmentally efficient manner.[8] Nevertheless, with the widespread usage of generative AI systems, these concerns equally emerged as societal concerns, in the sense that societies are currently struggling over how generative AI is and should be developed, deployed, and regulated. In other words, conflicts, negotiations, and power struggles are unfolding, which makes the need for a social science perspective beyond mere engineering concerns on how to live better with LLMs very apparent.

Given that the mentioned struggles exemplify how individual and collective actors try to secure autonomy within the development and deployment of generative AI, this article introduces the concept of digital self-determination as an analytical frame to understand, analyze, and conceptualize struggles over generative AI. An infrastructural perspective is presented as a focus point for an analysis of self-determination in relation to generative AI, as currently, many attempts to exercise self-determination in relation to generative AI are exercised on its infrastructures. Thus, to outline a perspective on generative AI as the object—and not the instrument—of obtaining or limiting autonomy, this article first presents a systematization of current literature on matters of sovereignty, autonomy, and self-determination in relation to digital media technologies. It then combines insights from related disciplines, for example, machine learning, to discuss how matters of autonomy relate to generative AI. These insights bring autonomy struggles into focus, which are performed in relation to generative AI infrastructures. The article concludes with a proposition for a research framework based on matters of digital self-determination that allows for multi-actor, process-oriented, and situated

[4] Bender et al., "Dangers of Stochastic Parrots," 613–15; Solon Barocas et al., "Fairness and Machine Learning": 171–201; Emma Strubell et al., "Energy and Policy Considerations for Deep Learning in NLP," preprint, arXiv, 2019, https://arxiv.org/abs/1906.02243.

[5] Irene Solaiman et al., "Evaluating the Social Impact."

[6] Timnit Gebru et al., "Datasheets for Datasets," preprint, arXiv, December 1, 2021, http://arxiv.org/abs/1803.09010.

[7] Shira Mitchell et al., "Algorithmic Fairness: Choices, Assumptions, and Definitions," *Annual Review of Statistics and Its Application* 8, no. 1 (March, 2021): 141–63, https://doi.org/10.1146/annurev-statistics-042720-125902.

[8] Gaurav Menghani, "Efficient Deep Learning: A Survey on Making Deep Learning Models Smaller, Faster, and Better," preprint, arXiv, June 21, 2021, http://arxiv.org/abs/2106.08962.

contextual analyses of how autonomy implications manifest in generative AI infrastructures.

Digitization as an object or instrument of autonomy

Matters of sovereignty, autonomy, and self-determination are diversely discussed within media and communication research. The three terms—digital sovereignty, digital autonomy, and digital self-determination—are often used both interchangeably and contradictorily.[9] The different approaches within this realm are loosely united by the idea and ability of self-governance, exercising free will, and the absence of determination or control by others in one's individual or collective interaction with media technologies.[10]

The term *sovereignty* has traditionally highlighted the centralization of power with a state sovereign and their ability to authoritatively act within a state territory.[11] In relation to digital media technologies, this usually implies a state's ability (or a challenge to its ability) to exercise agency in relation to global digital infrastructures. This is especially triggered by the intergovernmental organization of the internet and associated technologies and the increasing dependence on digital technology providers. However, the term has been increasingly applied beyond the state level as a matter of individual sovereignty, often in response to some form of hegemonic power, exercised on the state level (e.g., individuals vs. autocratic systems) or by technology providers (e.g., individuals vs. big tech).[12]

Even though no clear-cut distinction can be made, the term autonomy seems to have been applied more consistently at the individual level—often focusing on matters of manipulation and control in relation to user agency in digital media technologies and algorithmic systems.[13] Sahebi and Formosa, with reference to philosophical discourses on autonomy, further highlight the building of autonomy

---

[9] Stephane Couture and Sophie Toupin, "What Does the Notion of 'Sovereignty' Mean When Referring to the Digital?," *New Media & Society* 21, no. 10 (2019): 2305–22, https://doi.org/10.1177/1461444819865984.

[10] Couture and Toupin, "Notion of 'Sovereignty,'" 2317–18; Siavosh Sahebi and Paul Formosa, "Social Media and Its Negative Impacts on Autonomy," *Philosophy & Technology* 35, no. 3 (September 2022): 70, https://doi.org/10.1007/s13347-022-00567-7; Julia Pohle and Thorsten Thiel, "Digital Sovereignty," *Internet Policy Review* 9, no. 4 (December, 2020), https://doi.org/10.14763/2020.4.1532.

[11] Pohle and Thiel, "Digital Sovereignty"; Couture and Toupin, "Notion of 'Sovereignty'."

[12] Couture and Toupin, "Notion of 'Sovereignty'."

[13] Sahebi and Formosa, "Social Media"; Laura Savolainen and Minna Ruckenstein, "Dimensions of Autonomy in Human–Algorithm Relations," *New Media & Society* 26, no. 6 (2024): 3472–90, https://doi.org/10.1177/14614448221100802.

competencies as a crucial element in autonomy endeavors and any attempts to harm such competencies as threatening an individual's autonomy.[14] They additionally differentiate between global and situated autonomy—the former referring to a person's ability to have control over their life course and the latter highlighting specific and possibly singular situations in which autonomy may be harmed or exercised.

To avoid misunderstandings and terminological inconsistencies, I focus on autonomy as an umbrella term to incorporate the literature dealing with matters of self-governance, free will, and lack of determination by others, referring to the Greek "αὐτονομία" (autonomia) as "self-law" (autos – self/nomos – law). First, I present a structuring overview of the current literature in this field and then specify the analytical concept of self-determination to analyze current struggles in relation to generative AI.

Current literature in the broad field of autonomy and digitization can be roughly divided into 1) digitization as an object of autonomy and 2) digitization as an instrument within aspirations for autonomy. As the current literature on media and communication exemplifies, individual or collective autonomy can then either be seen as being limited or supported through digitization.

Research on *digitization as an object of autonomy* deals with questions of how individual and collective actors can self-govern in relation to the design, production, governance, and usage of digital media technologies. In other words, this field of research addresses how people and collectives can exercise free will in relation to digital media technologies and not be determined by others (e.g., platform companies) in their relation to media technologies.

Considering social networking sites, matters of external determination have been prevalent since their emergence. The types of usage a platform affords[15] are largely and very restrictively determined by platform companies. One early example is Mark Zuckerberg's seminal statement "You have one identity,"[16] forcing real-name policies upon users, thereby discriminating and possibly endangering, for instance, transgender people and activists.[17] Such determination and especially the mostly

[14] Sahebi and Formosa, "Social Media."

[15] Ian Hutchby, "Technologies, Texts and Affordances," *Sociology* 35, no. 2 (May 2001): 441–56, https://doi.org/10.1177/S0038038501000219.

[16] José Van Dijck, "'You Have One Identity': Performing the Self on Facebook and LinkedIn," *Media, Culture & Society* 35, no. 2 (March 2013): 199–215, https://doi.org/10.1177/0163443712468605.

[17] Oliver L. Haimson and Anna Lauren Hoffmann, "Constructing and Enforcing 'Authentic' Identity Online: Facebook, Real Names, and Non-Normative Identities," *First Monday* (June 2016), https://doi.org/10.5210/fm.v21i6.6791.

unsuccessful attempts by affected individuals and civil-society groups[18] thus far to have Facebook change its policy demonstrates the limits of self-governance when using digital media platforms.

In this context, individual as well as collective attempts to counter extensive data collection practices, privacy infringements, and the lack of customizations in digital platforms,[19] for instance, as part of the digital rights movement, describe attempts to exercise free will in relation to digital media practices in an environment designed to limit individual and collective autonomy. Nevertheless, especially regarding questions of algorithmic manipulation, research has equally highlighted the ability of individual users to resist algorithmic control, to *game the algorithm*, and to actively appropriate media technologies in divergent ways within algorithmic systems.[20]

Research on *digitization as an instrument of autonomy* views digital media technologies as assistance in obtaining or limiting individual or collective autonomy in a field different than the design, production, or usage of digital media. In contemporary research, autonomy may be conceptualized as a desirable goal in autocratic political systems and researched through the lens of how digital media may facilitate attempts to reach political autonomy or as a tool for surveillance to hinder attempts to gain autonomy.[21] In this context, internet shutdowns are described as a

---

[18] Eva Galperin and Wafa Ben Hassine, "Changes to Facebook's 'Real Names' Policy Still Don't Fix the Problem," Electronic Frontier Foundation, December 18, 2015, https://www.eff.org/deeplinks/2015/12/changes-facebooks-real-names-policy-still-dont-fix-problem.

[19] Helen Nissenbaum, "A Contextual Approach to Privacy Online," *Daedalus* 140, no. 4 (October 2011): 32–48, https://doi.org/10.1162/DAED_a_00113; Sabine Trepte, "The Social Media Privacy Model: Privacy and Communication in the Light of Social Media Affordances," *Communication Theory* 31, no. 4 (November 2021): 549–70, https://doi.org/10.1093/ct/qtz035; Paola Pierri and Elizabeth Calderon Lüning, "A Democratic Approach to Digital Rights: Comparing Perspectives on Digital Sovereignty on the City Level," *International Journal of Communication* (2023): 3600–18; Sebastian Sevignani, *Privacy and Capitalism in the Age of Social Media*, 1st ed. (Routledge, 2015), https://doi.org/10.4324/9781315674841; Marina Micheli et al., "Emerging Models of Data Governance in the Age of Datafication," *Big Data & Society* 7, no. 2 (2020), https://doi.org/10.1177/2053951720948087.

[20] Savolainen and Ruckenstein, "Dimensions of Autonomy"; Caitlin Petre et al., "'Gaming the System': Platform Paternalism and the Politics of Algorithmic Visibility," *Social Media + Society* 5, no. 4 (October 2019), https://doi.org/10.1177/2056305119879995.

[21] Hanan Badr and Lena-Maria Möller, "Beyond Mainstream Media and Communication Perspectives on the Arab Uprisings," *Media and Communication* 9, no. 4 (December 2021): 260–63, https://doi.org/10.17645/mac.v9i4.5151; Aim Sinpeng, "Digital Media, Political Authoritarianism, and Internet Controls in Southeast Asia," *Media, Culture & Society* 42, no. 1 (January 2020): 25–39, https://doi.org/10.1177/0163443719884052; Marlies Glasius and Marcus Michaelsen, "Authoritarian Practices in the Digital Age| Illiberal and Authoritarian Practices in the Digital Sphere—Prologue,"

political instrument to limit individual and collective autonomy,[22] exemplifying the relevance of digital media as an infrastructure of political action.

Additionally, in entertainment or commercial contexts, current media and communication literature describes how digital, and especially algorithmic media, limit the autonomy of individuals consuming cultural products or commercial goods and services online.[23] One key term in this context is manipulation,[24] as an exemplification of a lack of self-determination—a discussion that has gained much traction in recent debates about political opinion-making. Even before the case of Cambridge Analytica has the extensive online profiling of people for political targeting, or the influence of algorithmically curated news feeds been discussed in relation to manipulating people's opinion-making and electoral decisions.[25]

On the state level, matters of autonomy, or in this context more often labeled sovereignty, also concern the dependence of government services and operations on certain digital infrastructure providers, which is increasingly being considered to

---

International Journal of Communication 12 (2018): 2018, https://ijoc.org/index.php/ijoc/article/view/8899/2459; Yury Kabanov and Mikhail Karyagin, "Data-Driven Authoritarianism: Non-Democracies and Big Data," in *Digital Transformation and Global Society*, Communications in Computer and Information Science, ed. D. A. Alexandrov et al. (Springer International Publishing, 2018), 144–55, https://doi.org/10.1007/978-3-030-02843-5_12.

[22] Víctor Sampedro et al., "Digital Disintermediation, Technical and National Sovereignty: The Internet Shutdown of Catalonia's 'Independence Referendum,'" *European Journal of Communication* 37, no. 2 (2022): 127–44, https://doi.org/10.1177/02673231211012143.

[23] Savolainen and Ruckenstein, "Dimensions of Autonomy"; Ana Rita Gonçalves et al., "Artificial Intelligence vs. Autonomous Decision-Making in Streaming Platforms: A Mixed-Method Approach," *International Journal of Information Management* 76 (2024), https://doi.org/10.1016/j.ijinfomgt.2023.102748; Jonathan Gingerich, "Is Spotify Bad for Democracy? Artificial Intelligence, Cultural Democracy, and Law," *Yale Journal of Law and Technology* 24 (2022): 227–316.

[24] Daniel Susser et al., "Technology, Autonomy, and Manipulation," *Internet Policy Review* 8, no. 2 (June 30, 2019), https://doi.org/10.14763/2019.2.1410.

[25] Andrew M. Guess et al., "How Do Social Media Feed Algorithms Affect Attitudes and Behavior in an Election Campaign?," *Science* 381, no. 6656 (July 2023): 398–404, https://doi.org/10.1126/science.abp9364; Yochai Benkler et al., *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (Oxford University Press, 2018), https://doi.org/10.1093/oso/9780190923624.001.0001; Emilio Ferrara et al., "Characterizing Social Media Manipulation in the 2020 U.S. Presidential Election," *First Monday* (October 2020), https://doi.org/10.5210/fm.v25i11.11431; Jacquelyn Burkell and Priscilla M. Regan, "Voter Preferences, Voter Manipulation, Voter Analytics: Policy Options for Less Surveillance and More Autonomy," *Internet Policy Review* 8, no. 4 (December 2019), https://doi.org/10.14763/2019.4.1438.

limit a state's sovereignty, especially in times of increasing geopolitical tensions.[26] In some countries, this has led to a complete ban of certain online platforms.[27] Furthermore, this is currently reflected in the attempts by the Biden administration in the US to enforce legislation that would either ban TikTok from the US market or force its parent company ByteDance to sell the platform.[28]

The two dimensions of digitization—as an *object* or an *instrument* of autonomy—are closely interrelated. The massive collection of personal data as well as the possibility for personalized targeting is one of the many reasons individuals, digital rights activists, and policymakers demand changes to the way social media platforms operate. Famously, Shoshana Zuboff has argued in this context that surveillance capitalism, exemplified through major online platforms, undermines individual and collective autonomy overall:

> The competition for surveillance revenues bears down on our bodies, our automobiles, our homes, and our cities, challenging human autonomy and democratic sovereignty in a battle for power and profit as violent as any the world has seen.[29]

Broken down to everyday lived experiences of individuals, this lack of self-governance in surveillance capitalism describes how the metrification and datafication of social interactions, the extensive profiling of individuals through their digital traces, and the invasive targeted advertising systems and recommender algorithms limit people's freedom of will and self-governance.

Although some of the outlined implications also apply to generative AI, new autonomy implications are emerging. In the following sections, I specifically focus on generative AI as the object of autonomy by demonstrating how current struggles over generative AI target the technology itself. These struggles exemplify that individuals and collectives try to act on[30] generative AI as a technology to change its current

---

[26] Bernardus Jansen et al., "Pushing Boundaries: An Empirical View on the Digital Sovereignty of Six Governments in the Midst of Geopolitical Tensions," *Government Information Quarterly* 40, no. 4 (2023), https://doi.org/10.1016/j.giq.2023.101862.

[27] Anilesh Kumar and Daya Thussu, "Media, Digital Sovereignty and Geopolitics: The Case of the TikTok Ban in India," *Media, Culture & Society* 45, no. 8 (November 2023): 1583–99, https://doi.org/10.1177/01634437231174351.

[28] Bobby Allyn, "President Biden signs law to ban TikTok nationwide unless it is sold," *NPR*, April 24, 2024, https://www.npr.org/2024/04/24/1246663779/biden-ban-tiktok-us.

[29] Shoshana Zuboff, "Surveillance Capitalism and the Challenge of Collective Action," *New Labor Forum* 28, no. 1 (January 2019): 11, https://doi.org/10.1177/1095796018819461.

[30] Sigrid Kannengießer and Sebastian Kubitschko, "Acting on Media: Influencing, Shaping and (Re)Configuring the Fabric of Everyday Life," *Media and Communication* 5, no. 3 (2017): 1–4, https://doi.org/10.17645/mac.v5i3.1165.

course—rather than to act with generative AI to obtain autonomy. Focusing on generative AI as the object of autonomy struggles make the need for an infrastructural perspective apparent.

Autonomy in relation to generative AI

Generative AI refers to text- or image-generating systems that have become widely popular with ChatGPT as well as with similar systems, such as Dall·E, Gemini, Claude, and Stable Diffusion. Current generative AI systems are mostly built upon LLMs, which might become integrated into multimodal models, where systems are able to process different modalities of data for instance in text-to-image generation. These models are very capable in string-prediction tasks. In other words, within a sequence of words (or other tokens), they predict further sequence elements to produce, for example, coherent-sounding text. In attempts to describe what LLMs do, some have described them as gigantic and much more complex autocomplete systems. Critical scholars have called them "stochastic parrots,"[31] because these models largely base their string predictions on probability, and similarly to parrots who imitate speech, they do not have access to meaning.

LLMs have not been invented with the launch of ChatGPT. The first transformer model architecture, which presented the technological breakthrough that makes current LLMs so capable, was presented in 2017, with the first BERT model being introduced by Google in 2018.[32] The novelty of ChatGPT was that in 2022, OpenAI connected the LLM GPT-3, originally developed in 2020, to a user interface that made the model accessible for layperson usage. This widespread application of LLMs through generative AI systems has led to struggles inherent to the LLM architecture becoming center stage, which already had a trajectory in such different academic fields as AI ethics, critical machine learning, computational linguistics, and critical data and media studies.

Based on systematization of the literature from these fields, the following sections points towards *training data*, *accountability*, *market power,* and *social and environmental justice implications* to selectively specify autonomy struggles regarding generative AI. These examples relate to matters auf autonomy since they demonstrate a lack of control over the use of personal data or copyright (training data), obstacles

[31] Emily M. Bender et al., "On the Dangers of Stochastic Parrots."

[32] Ashish Vaswani et al., "Attention Is All You Need," preprint, arXiv, June 2, 2017, https://doi.org/10.48550/ARXIV.1706.03762; Jacob Devlin et al., "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding," preprint, arXiv, May 24, 2019, http://arxiv.org/abs/1810.04805.

in attempts to self-govern in relation to generative AI (accountability), determination in the application of generative AI by few dominant actors (market concentration) and the endangering of self-determined ways of life (social and environmental justice). They demonstrate the need for an analytical frame that conceptualizes generative AI as the object of autonomy, especially along its infrastructures, where individuals and collectives act on the materiality of generative AI[33] to secure autonomy.

Such an object-oriented perspective has often been overlooked in media and communication research. However, as a part of an infrastructural turn in the field, Lisa Parks and Nicole Starosielski have poignantly described how an infrastructural perspective addresses the deliberate invisibility of media infrastructures and their profound political implications. An infrastructural perspective on generative AI thus provides an important addition to contemporary approaches toward generative AI in media and communication research, which has thus far focused largely on matters of symbolic interaction[34]. It focuses on the power structures and justice implications—and thus on the far-reaching societal implications that accompany the shift toward generative AI.

Autonomy in relation to training data

LLMs rely on large datasets for their training, which are typically scraped from online sources. Although datasets are available from nonprofit organizations,[35] especially large generative AI companies increasingly rely on their independently collected datasets.[36] These companies usually do not provide information on the constitution of their datasets, which has triggered suspicions that they rely on copyrighted material. Consequently, several copyright holders, most prominently the New York Times, several artists and major record labels including Sony Music, Warner Music Group, and Universal Music Group, have sued generative AI companies over copyright

[33] Anne Mollen and Sigrid Kannengießer, "Shaping AI (more) sustainably. Socio-technological perspectives on AI infrastructures," in *Uncertain journeys into digital futures,* ed. Thomas Kox, André Ullrich, and Herbert Zech (Nomos, 2025).

[34] Andrea L. Guzman and Seth C. Lewis, "Artificial Intelligence and Communication: A Human–Machine Communication Research Agenda," *New Media & Society* 22, no. 1 (2020): 70–86, https://doi.org/10.1177/1461444819858691.

[35] Stefan Baack, "Training Data for the Price of a Sandwich," *Mozilla Foundation,* February 6, 2024, https://foundation.mozilla.org/en/research/library/generative-ai-training-data/common-crawl/.

[36] Kevin Schaul et al., "Inside the secret list of websites that make AI like ChatGPT sound smart," *Washington Post*, April 19, 2023, https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning/.

violations.[37] In attempts to secure their rights against generative AI providers, they argue that the use of their material for training purposes does not fall under fair-use scenarios. These copyright violations are relevant from an autonomy perspective as copyright holders are trying to exercise control over their material and not be limited in their rights by generative AI providers.

A similar example can be found at the level of individual users. The collection of training data from interpersonal communication has recently led to fears over severe privacy violations by social media companies. For example, the Austrian civil society organization NOYB has filed complaints against Meta with 11 Data Protection Authorities in Europe to have Meta stop using private communication for AI systems training. As a result, Meta has presumably stopped using EU citizens' data for such purposes.[38] Here, users and their representatives fear a loss of control and determination through social media companies in their handling of personal data for training purposes.

Responses in engineering-oriented fields have addressed matters of autonomy over training data primarily through the lens of transparency and quality of training data. In attempts to make generative AI systems better and more aligned with data protection matters, the field has proposed data sheets as transparency tools, the careful curation of smaller datasets, or systems that unlearn personal data.[39] Although such technical approaches assist individuals in reaching more autonomy on a system level, many struggles over autonomy in relation to training data do not primarily

---

[37] The New York Times, "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work," *New York Times*, December 27, 2023, https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html; Emilia David, "Sarah Silverman's lawsuit against OpenAI partially dismissed," *The Verge*, February 13, 2024, https://www.theverge.com/2024/2/13/24072131/sarah-silverman-paul-tremblay-openai-chatgpt-copyright-lawsuit; James O'Donnell, "Training AI music models is about to get very expensive," *The Verge*, June 27, 2024, https://www.technologyreview.com/2024/06/27/1094379/ai-music-suno-udio-lawsuit-record-labels-youtube-licensing/?truid=&utm_source=the_algorithm&utm_medium=email&utm_campaign=the_algorithm.unpaid.engagement&utm_content=07-08-2024&mc_cid=982c1b8019.

[38] noyb, "(Preliminary) noyb WIN: Meta stops AI plans in the EU," press release, *noyb*, 14 June, 2024, https://noyb.eu/en/preliminary-noyb-win-meta-stops-ai-plans-eu.

[39] Gebru et al., "Datasheets for Datasets"; Orestis Papakyriakopoulos et al., "Augmented Datasheets for Speech Datasets and Ethical Decision-Making," in *2023 ACM Conference on Fairness, Accountability, and Transparency, FAccT '23* (ACM, 2023): 881–904, https://doi.org/10.1145/3593013.3594049; Menghani, "Efficient Deep Learning"; Sebastian Schelter et al., "Forget Me Now: Fast and Exact Unlearning in Neighborhood-Based Recommendation," in *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '23* (ACM, 2023): 2011–15, https://doi.org/10.1145/3539618.3591989.

concern the technical system. Instead, they relate to organizational practices of misusing private and copyrighted data, the market power of big-tech companies, and a perceived ability to dictate new industry practices of how to use data, even if these may not align with legal frameworks. Additionally, these struggles have a discursive dimension because legitimation strategies on what constitutes fair use and legitimate interests are conducted discursively, exemplified by OpenAI, which claimed in January 2024 that it is essentially impossible to train generative AI without copyrighted material.[40] Thus, a multitude of possible social science-oriented research questions emerge on matters of autonomy and training data of generative AI.

Autonomy in relation to accountability

Since the release of generative AI applications, several problematic tendencies and social impacts in generated outputs have become apparent:[41] cultural misrepresentations, discriminatory outputs, the use within targeted disinformation campaigns, and factually incorrect or privacy-violating outputs. Accountability is generally conceived as a solution to mitigate such harms by holding providers of generative AI responsible for their outputs. These harms exemplify the ways in which generative AI impacts people's ability to exercise control over their own lives, for instance, by being adequately represented or not having their rights violated. Although generative AI companies increasingly recognize the issue and work toward technical solutions, the immanent lack of accountability on behalf of generative AI providers continues to violate people's ability to self-govern in different ways.

Recent examples include defamation cases against major tech companies, for instance in cases where generative AI was used to aggregate news articles leading to misinformation. Misinformation through generative AI about individuals might severely impact people's livelihoods. A recent court case has accordingly been described as "a new frontier in libel law,"[42] indicating that currently, matters of

[40] Benj Edwards, "OpenAI says it's 'impossible' to create useful AI models without copyrighted material," *Ars Technica*, January 9, 2024, https://arstechnica.com/information-technology/2024/01/openai-says-its-impossible-to-create-useful-ai-models-without-copyrighted-material/.

[41] Solaiman et al., "Evaluating the Social Impact"; Pegah Maham and Sabrina Küpert, "Governing General Purpose AI: A Comprehensive Map of Unreliability, Misuse and Systemic Risks," Policy Brief (Stiftung Neue Verantwortung, 2023), https://www.stiftung-nv.de/sites/default/files/snv_governing_general_purpose_ai_pdf.pdf.

[42] Mary Carolan, "RTÉ's Dave Fanning initiates defamation case after his photo appeared in article about sexual misconduct trial of different broadcaster," *The Irish Times*, January 15, 2024,

accountability regarding generative AI outputs are not well-suited to protect people's autonomy. These so far still isolated cases exemplify an underlying systematic need to secure accountability as a precondition for self-governance in relation to generative AI. This becomes all the more clear, in the case of targeted disinformation campaigns, which are feared to operate at scale through the use of generative AI,[43] resulting in widespread political misinformation, which may seriously endanger social cohesion and democratic systems. Generative AI companies have started to address their role in political decision-making;[44] however, they do not offer convincing solutions or even basic information[45] when challenged about factually incorrect outputs about ongoing elections.[46] In particular, European and other political bodies are currently attempting to establish practices that will hold generative AI providers accountable for their role in political disinformation[47] in attempts to exercise control and establish adequate governance framework to secure individual autonomy and collective sovereignty.

Matters of accountability in relation to AI outputs are a long-established concern[48], especially as response to the severe impacts as outlined above that the technology can have on people's ability to exercise control over their own lives. Considering the difficulty to hold generative AI providers accountable for violations of people's and collectives' autonomy, engineering-oriented fields and policy-making have proposed more transparency, for instance information on architecture, model size, training data, computational resources needed, safety measures, and testing strategies. Nontransparency is indeed an issue in AI since it undermines accountability understood as a relational process of answerability – where information about and justification of one's actions are key for ensuring

https://www.irishtimes.com/crime-law/courts/2024/01/15/rtes-dave-fanning-initiates-defamation-case-after-his-photo-appeared-in-article-about-sexual-misconduct-trial-of-different-broadcaster/.

[43] Johannes Sedlmeir et al., "Battling Disinformation with Cryptography," *Nature Machine Intelligence* 5, no. 10 (October 2023): 1056–57, https://doi.org/10.1038/s42256-023-00733-2.

[44] OpenAI, "How OpenAI is approaching 2024 worldwide elections," *OpenAI* blog, January 15, 2024, https://openai.com/index/how-openai-is-approaching-2024-worldwide-elections/.

[45] Eliza Grkitsi, "Commission 'compels' Microsoft to provide information on Bing and generative AI," *Euractiv*, March 17, 2024, https://www.euractiv.com/section/digital/news/commission-compels-microsoft-to-provide-information-on-bing-and-generative-ai/.

[46] Clara Helming, "AI Chatbot produces misinformation about elections," *AlgorithmWatch*, December 15, 2023, https://algorithmwatch.org/en/study-microsofts-bing-chat/.

[47] European Commission, "Code of Practice on Disinformation: new reports available in the Transparency Centre," press release, September 26, 2023, https://digital-strategy.ec.europa.eu/en/news/code-practice-disinformation-new-reports-available-transparency-centre.

[48] Claudio Novelli et al., "Accountability in artificial intelligence: what it is and how it works," *AI & Society* 39 (2024): 1871–82, https://doi.org/10.1007/s00146-023-01635-y.

accountability.[49] However, beyond enforcing and enabling transparency through engineering and policy solutions, there is much to be investigated about how autonomy is specifically challenged or ensured through accountability. For instance, questions of accountability are also played out discursively, where it needs to be investigated who is to be held accountable for what exactly. What are indeed the autonomy implications resulting from a lack of accountability regarding generative AI—do they lie primarily within the technical system or are they also rooted in the appropriation of generative AI within everyday practices of use? Accountability and autonomy thus do not only relate to the technical system itself and how to better engineer and regulate such systems, but also concern people's and collectives' attempts to shape accountability to ensure autonomy.

Autonomy in relation to market concentration

Questions of autonomy imply matters of determination by others. In relation to generative AI, this necessarily leads to questions about the different types of systems from which users can choose. From a market perspective, it has been postulated that the generative AI industry is already heavily concentrated and is from its outset prone to concentration.[50] Accordingly, US and European competition authorities, as well as various national ones, have launched investigations into different mergers, acquisitions, and cooperations in relation to generative AI.

Network effects are one reason the generative AI industry may be particularly prone to concentration. Other reasons include the reliance on a workforce with the knowledge necessary for building LLMs, access to massive amounts of computational resources needed to develop and train LLMs, and access to gigantic datasets to train LLMs. These preconditions establish very high entry barriers to the generative AI market, giving the already dominant tech companies an advantage in the generative AI race. As a result, a select few companies dominate the generative AI market.[51] In

[49] Claudio Novelli et al., "Accountability in artificial intelligence," 1872.

[50] Jai Vipra and Anton Korinek, "Market Concentration Implications of Foundation Models: The Invisible Hand of ChatGPT," Center on Regulation and Markets Working Paper #9 (Brookings, 2023), https://www.brookings.edu/wp-content/uploads/2023/09/market-concentration-implications-of-foundation-models-final-1.pdf; FTC, "Generative AI Raises Competition Concerns," Federal Trade Commission, 2023, accessed March 3, 2025,https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns.

[51] Mark Surman et al., "Accelerating Progress Toward Trustworthy AI," *Mozilla Foundation*, February 22, 2024, https://foundation.mozilla.org/en/research/library/accelerating-progress-toward-trustworthy-ai/whitepaper/.

addition, often-formulated calls for open generative AI should be considered cautiously. Experts around Meredith Whitaker, the president of the Signal Foundation, have argued that although openness can leverage transparency, accountability, and accessibility to build applications on top of existing generative AI models,

> some companies have moved to embrace "open" AI as a mechanism to entrench dominance, using the rhetoric of "open" AI to expand market power, and investing in "open" AI efforts in ways that allow them to set standards of development while benefiting from the free labor of open source contributors.[52]

This combination of factors makes it very difficult for alternative generative AI systems to be built, which leaves little flexibility for users to choose alternatives to the dominant products available on the market. Beyond diversity in the product range, dominant generative AI providers are currently setting the norms on how to produce LLMs, establishing mostly extractive production practices for LLMs (see the next section). This has been one reason the Mozilla Foundation has been calling for changes in AI development norms and started to test and experiment with alternative generative AI developments.[53]

However, the concentration in the generative AI market does not concern only individual autonomy in choosing systems that do not violate one's personal principles and values. They also have led to state-level, and thus sovereignty, concerns about the dominance of US products. As a response, European countries and the European Union have started investment initiatives to assemble facilities that can provide the necessary compute for developing, training, and running European LLMs. Furthermore, they are looking to make these resources accessible to smaller and medium-sized enterprises.[54] Moreover, commission president Ursula von der Leyen made the building of these so-called AI factories a topic in her 2023 State of the Union address,[55] demonstrating that the European Commission assigns high priority to this struggle to autonomously develop and use European LLMs. Further, in relation to antitrust law, the European Commission launched an investigation into specific

[52] David Gray Widder et al., "Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI," *SSRN Electronic Journal* (2023), https://doi.org/10.2139/ssrn.4543807.

[53] Mark Surman et al., "Accelerating Progress."

[54] European Commission, "Commission Launches AI Innovation Package to Support Artificial Intelligence Startups and SMEs," news release, January 24, 2024, https://digital-strategy.ec.europa.eu/en/news/commission-launches-ai-innovation-package-support-artificial-intelligence-startups-and-smes.

[55] European Comission, State of the Union address, September 13, 2023, https://state-of-the-union.ec.europa.eu/state-union-2023_en.

partnerships between tech giants (e.g., Microsoft, Apple, and OpenAI)[56] as well as calls for contributions[57] on how well EU competition law is equipped to deal with market concentration in the generative AI market, demonstrating the need for reflecting on regulatory frameworks in changing market conditions to protect autonomy.

Autonomy in relation to social and environmental justice

Many struggles over data centers and their resource consumption, the resource extraction necessary for building hardware, the pollution that often comes with hardware disposal, and precarious working conditions along the entire generative AI value chain can be seen when the consequences of generative AI in the Global South are considered. Although digital infrastructures are not particular to generative AI but underly all digital applications,[58] the massive resource needs of LLMs justify considering it a specific site for struggle in relation to generative AI. After all, the large-scale deployment of LLMs has further aggravated environmental and social justice concerns over energy, water, and minerals needed to run data centers and manufacture computing hardware[59] as well as over precarious working conditions of the workers who are needed to curate data and generative AI outputs.[60]

Current examples demonstrate how activists and Indigenous communities, for instance, in drought-stricken regions, are fighting against the construction of hyperscale data centers with their immense water consumption, endangering the

[56] Julia Tar, "EU Commission to examine Microsoft-OpenAI partnership," Euractiv, accessed January 9, 2024, https://www.euractiv.com/section/competition/news/eu-commission-to-examine-microsoft-openai-partnership; Julia Tar, "Apple's ChatGPT integration raises data privacy and competition questions," *Euractiv*, June 17, 2024, https://www.euractiv.com/section/data-privacy/news/apples-chatgpt-integration-raises-data-privacy-and-competition-questions/.

[57] European Commission, "Commission launches calls for contributions on competition in virtual worlds and generative AI," press release, January 9, 2024, https://ec.europa.eu/commission/presscorner/detail/en/ip_24_85.

[58] Lisa Parks et al., *Media Backends Digital Infrastructures and Sociotechnical Relations* (University of Illinois Press, 2023); Lisa Parks and Nicole Starosielski, eds., *Signal Traffic: Critical Studies of Media Infrastructures* (University of Illinois Press, 2015).

[59] Alexandra Sasha Luccioni et al., "Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model," preprint, arXiv, November 3, 2022, https://arxiv.org/abs/2211.02001; Bender et al., "Dangers of Stochastic Parrots"; David Mytton, "Data Centre Water Consumption," *NPJ Clean Water* 4, no. 1 (2021): 11, https://doi.org/10.1038/s41545-021-00101-w; Pengfei Li et al., "Making AI Less 'Thirsty': Uncovering and Addressing the Secret Water Footprint of AI Models," preprint, arXiv, 2023, https://doi.org/10.48550/ARXIV.2304.03271; Emma Strubell et al., "Energy and Policy Considerations for Deep Learning in NLP."

[60] Gray and Siddharth, *Ghost Work*; Kate Crawford, *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence* (Yale University Press, 2021).

water supply of local communities. These forms of extraction of local territories continue colonial power imbalances and exploitative relations in the Global North as well, for instance in Ireland, which has a high concentration of big-tech data centers.[61] Other forms of extraction concern the precarious working conditions currently implemented along the LLM value chain, with workers specifically in the Global South starting to demand more adequate pay and psychological support structures for the often extremely challenging work on datasets and generative AI outputs.[62]

With the high market concentration in the generative AI industry described above, major generative AI providers have established these extractive norms in LLM production. As a result, the people affected by these extractivist practices have difficulties in gaining autonomy over their livelihoods and territories, which are to some extent controlled by generative AI providers. Moreover, users of generative AI are left with nearly no options when trying to choose generative AI products aligning with their personal environmental and social justice beliefs.

Self-determination as a perspective on generative AI infrastructures

Many of the discussed autonomy implications and associated conflicts are rooted in the automated systems themselves. Discrimination through generative AI lies, for example, in the programming of the systems, the datasets used, or the organizational use of the AI-generated outputs.[63] The ecological sustainability impacts, in turn, depend on the size of the models, the datasets and hardware used, the electricity mix of the data centers, and the organizational decisions on when and where the models

[61] Marcela Suárez Estrada and Sebastián Lehuedé, "Towards a Terrestrial Internet: Re-Imagining Digital Networks from the Ground Up," *Tapuya: Latin American Science, Technology and Society* 5, no. 1 (2022), https://doi.org/10.1080/25729861.2022.2139913; Patrick Bresnihan and Patrick Brodie, "New Extractive Frontiers in Ireland and the Moebius Strip of Wind/Data," *Environment and Planning E: Nature and Space* 4, no. 4 (2021): 1645–64, https://doi.org/10.1177/2514848620970121; Patrick Brodie, "Data Infrastructure Studies on an Unequal Planet," *Big Data & Society* 10, no. 1 (2023), https://doi.org/10.1177/20539517231182402.

[62] Billy Perrigo, "OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic," *Time Magazine*, January 18, 2023, https://time.com/6247678/openai-chatgpt-kenya-workers/; Uma Rani et al., "Development of AI: Role of "invisible workers" in the AI value chains," May 2024, Science-Policy Brief for the Multistakeholder Forum on Science, Technology and Innovation for the SDGs, May, 2024, https://sdgs.un.org/sites/default/files/2024-05/Rani%3B%20Gobel%3B%20Dhir_Development%20of%20AI.pdf.

[63] Barocas et al., "Fairness and Machine Learning."

are trained.[64] These examples illustrate that the extent to which automated systems can be harmonized with individual and collective autonomy efforts is not primarily negotiated on the user level, on which end users interact with a generative AI system through the user interface. Instead, an infrastructural perspective seems adequate, as it focuses on generative AI infrastructures and how they manifest certain autonomy implications. Therefore, I propose a conceptual specification toward digital self-determination to analytically grasp the interrelation of autonomy efforts and generative AI infrastructures.

The infrastructural turn in communication and media studies has brought infrastructures like data, data centers, and online platforms into focus.[65] Understood as socio-technical systems, the analysis of infrastructures reflects on the agency and societal implications associated with the materiality of media technologies. In this context, infrastructures and their analyses increasingly point to political dimensions in the constitution and manifestation of digital media far beyond their interfaces.[66] This political dimension is ultimately about negotiating and shaping infrastructures in the context of self-governance.[67]

In the traditional sense, media infrastructures include data centers, data transmission networks, transmission masts, underwater cables, end devices, and software. Conceptually, they can be understood as "situated sociotechnical systems that are designed and configured to support the distribution of audiovisual signal traffic."[68] These classic infrastructures are also relevant for generative AI. Algorithms and generative AI systems are developed, trained, and used daily on huge server systems in data centers, where data are transmitted via submarine cables and then stored. Simultaneously, infrastructures for generative AI can also be located at the software level—from LLMs with their complex architectures and the immense

[64] Bender et al., "Dangers of Stochastic Parrots"; Li et al., "Making AI Less 'Thirsty'"; Mytton, "Data Centre Water Consumption."

[65] Brodie, "Data Infrastructure Studies"; Patrick Brodie and Julia Velkova, "Cloud Ruins: Ericsson's Vaudreuil-Dorion Data Centre and Infrastructural Abandonment," *Information, Communication & Society* 24, no. 6 (2021): 869–85, https://doi.org/10.1080/1369118X.2021.1909099; Panos Constantinides et al., "Introduction—Platforms and Infrastructures in the Digital Age," *Information Systems Research* 29, no. 2 (2018): 381–400, https://doi.org/10.1287/isre.2018.0794; Jean-Christophe Plantin and Aswin Punathambekar, "Digital Media Infrastructures: Pipes, Platforms, and Politics," *Media, Culture & Society* 41, no. 2 (2019): 163–74, https://doi.org/10.1177/0163443718818376.

[66] Parks et al., *Media Backends*.

[67] Francesca Musiani, "*Infrastructuring* Digital Sovereignty: A Research Agenda for an Infrastructure-Based Sociology of Digital Self-Determination Practices," *Information, Communication & Society* 25, no. 6 (April 2022): 785–800, https://doi.org/10.1080/1369118X.2022.2049850.

[68] Parks and Starosielski, *Signal Traffic*.

datasets that are required for their training to the fine-tuning of LLMs down the application chain.[69]

Many of the mentioned autonomy struggles with generative AI as its object relate to an infrastructural dimension. Negotiating generative AI infrastructures in reaching more autonomy could relate to, for instance, designing LLM architectures in a way that minimizes their environmental impact, the curation of datasets that do not produce discriminatory outputs and are not based on the exploitation of clickworkers in the Global South or on data transfer hubs that effect geopolitical shifts in power. An infrastructural perspective on generative AI infrastructures can accordingly be used to highlight how infrastructures come into being, how they may be changed, or how they manifest certain power imbalances and "arrange and materialize underlying relationships."[70] To analyze such practices of infrastructuring[71] in the context of autonomy discussions, I introduce the concept of *digital self-determination* to specify an analytical entry point toward the multiple attempts to self-govern in relation to generative AI.

This perspective presents a demarcation from autonomy and sovereignty perspectives in several respects. First, it does not focus on the concentration of power within a state-centric understanding (e.g., sovereignty discourse) but considers very diverse actors as essential for understanding autonomy struggles over generative AI. Second, it does not consider autonomy primarily as an objective when dealing with generative AI but highlights that self-determination is negotiated, contested, and established along generative AI infrastructures over time. Third, it clarifies that the ways infrastructures are being produced and changed occur in very specific situations, contexts and power constellations, thus, in other words, through situated practices. This provides analytical entry points to the study of self-determination in relation to generative AI through a practice theoretical lens. In other words, relying on self-determination as a specification in the autonomy discourse allows for analyses that consider a *multi-actor*, *process,* and *situated contextual* perspective.

Referring to the autonomy struggles in relation to generative AI as outlined above, such an approach would allow to empirically grasp how different actors try to exercise self-determination by gaining control over their personal data or copyrights, when being confronted with the current training data practices by generative AI

---

[69] Anne Mollen et al., *AI Infrastructures and Sustainability: Expanding Perspectives on Automation, Communication and Media* (Palgrave Macmillan, 2025).

[70] Musiani, "*Infrastructuring* Digital Sovereignty."

[71] Volkmar Pipek and Volker Wulf, "Infrastructuring: Toward an Integrated Perspective on the Design and Use of Information Technology," *Journal of the Association for Information Systems* 10, no. 5 (2009): 306–32.

providers that violate privacy or copyrights. It could also be used to analyze how attempts to self-govern in relation to generative AI requires adequate accountability frameworks to identify the levers how to self-govern and shape generative AI according to different norms and values. Considering the determination by a handful of major generative AI providers in a heavily concentrated market, a perspective on self-determination in relation to generative AI infrastructures helps in understanding the effects of determination and the barriers in gaining autonomy. Lastly, by focusing on social and environmental justice implications of generative AI infrastructures, it becomes also possible to consider much wider autonomy implications and related struggles that do not concern the application of generative AI directly, but rather how its infrastructures impact on livelihoods and territories along the generative AI value chain. A focus on self-determination along generative AI infrastructures then also allows to transfer situated empirical analyses beyond these individual cases. An infrastructure approach theorizes how self-determination might manifest or fail to do so over time in the materiality of infrastructures and their respective actor and power constellations.

Analyzing generative AI infrastructures from a perspective of self-determination broadens the view toward the manifold negotiation processes of generative AI beyond the direct human–technology interface. In this process, *multiple actors* become relevant in their autonomy struggles in relation to generative AI: from Indigenous populations who see themselves trapped in neocolonial dependencies in the face of expanding data infrastructures[72] to private households who have to reckon with uncertainties in the power supply in view of the establishment of large data centers[73] or to clickworkers who process data for the training of AI models and whose working conditions are reflected in the quality of the data[74]. In their autonomy struggles, all these actors play a role in constituting the norms through which generative AI and its infrastructures are coming into being and in shaping how generative AI is being implemented within societies. Consequently, from a self-determination perspective, it is crucial to analyze their struggles—to understand how their individual ability for self-governance may be harmed through generative AI

[72] Sebastian Lehuedé, "Territories of data: ontological divergences in the growth of data infrastructure," *Tapuya: Latin American Science, Technology and Society* 5, no. 1 (2022), https://doi.org/10.1080/25729861.2022.2035936.

[73] Frans Libertson et al., "Data-center infrastructure and energy gentrification: perspectives from Sweden," *Sustainability: Science, Practice and Policy* 17, no. 1 (2021): 152–61, https://doi.org/10.1080/15487733.2021.1901428,

[74] Milagros Miceli et al. "Documenting Data Production Processes: A Participatory Approach for Data Work," Proc. ACM Hum.-Comput. Interact. 6, CSCW2, Article 510, July 11, 2022, https://doi.org/10.48550/arXiv.2207.04958.

production and deployment processes and to realize beyond the individual level in what ways and through which actors generative AI manifests itself within societies.

A self-determination perspective equally focuses on the *processes* of struggling over self-governance in relation to generative AI, thereby acknowledging how certain norms with their respective autonomy implications become established and change over time. Taking struggles over the training data for LLMs as an example, it can be currently seen how different actors negotiate what data can legitimately be used for generative AI development. Different actors are attempting to establish certain practices of using copyrighted material as new norms, although they may violate current legislation. By contrast, other actors are trying to oppose the establishment of such norms. Analyzing how this process leads to certain norms becoming established can be very insightful for understanding how infrastructures of generative AI (e.g., datasets) come into being and what types of autonomy violations they may entail. Thus, a self-determination perspective not only analyses how strategies of sovereignty and autonomy might be inscribed and enacted within digital infrastructures[75] but also how the idea of autonomy itself becomes negotiated and manifested in these processes.

Negotiations on autonomy in relation to generative AI occur in courtrooms, data center sites, the offices of generative AI developers, and many other places. In that sense, generative AI infrastructures are not disconnected from the spaces of their manifestation but should be analyzed in their respective *situated contexts*. A self-determination analysis of generative AI thus requires empirical research that analyses how self-governance is enabled, undermined, or contested locally. It means that individual actors enact struggles over self-determination in specific settings and situations, which allow for detailed empirical analyses. This could be clickworkers negotiating working conditions in their Kenyan offices, Indigenous communities negotiating the data center technologies with big-tech companies along their respective resource consumption, or local communities engaging in alternative generative AI developments that acknowledge their cultural heritage.

Such analyses—multi-actor, process-oriented, and contextualized—will eventually help in understanding how certain practices of infrastructuring generative AI come into being, at a moment in time when these negotiations occur on a large scale. They will eventually help in investigating how impacts on autonomy manifest themselves in certain ways of "doing" generative AI. They allow for granular investigations into how this impact is constituted—in the negotiations between multiple actors in very diverse settings and over time—and how these newly

[75] Musiani, "*Infrastructuring* Digital Sovereignty."

establishing norms, routines and habits in relation to generative AI come to matter for individual and collective autonomy.

Conclusion: Living with generative AI

An autonomy-inspired perspective on generative AI exemplifies that its impact extends far beyond its outputs. It is thus crucial to recognize the influence of generative AI in media and communication research beyond matters of symbolic interaction. As research in critical machine learning has demonstrated, generative AI can have immense consequences for matters of individual and collective autonomy. Much can be learned from such discussions in media and communication research. However, although these insights from engineering disciplines are enlightening, they remain focused on the technical system itself. They recognize the impact of generative AI on people and societies;[76] nevertheless, it is not in the tradition of their field to address social processes involved in establishing such impacts. Media and communication research must move beyond research on generative AI that is devoid of any materiality and "object-less." Instead, it should acknowledge generative AI infrastructures as crucial sites of social action, where empirically, it can be investigated how the impacts on autonomy as self-governance manifest themselves.

The manifold struggles over generative AI this article describes provide empirically rich examples of how certain ways of doing generative AI become established, how they manifest in generative AI infrastructures, and how this affects matters of autonomy. Thus, the implications of generative AI on self-determination can be better understood through multi-actor, process-oriented, and contextual analyses of social practices in relation to generative AI infrastructures. Such analyses not only provide valuable insights into matters of social justice and power dynamics in societies increasingly shaped by generative AI but also build a solid ground for policy-making and transformative approaches toward generative AI, as they allow the investigation of how autonomy as a collective good may be impeded by how generative AI is developed, produced, and used.

[76] Solaiman et al., "Evaluating the Social Impact."

Bibliography

Allyn, Bobby. "President Biden signs law to ban TikTok nationwide unless it is sold." *NPR*, April 24, 2024, https://www.npr.org/2024/04/24/1246663779/biden-ban-tiktok-us.

Baack, Stefan. "Training Data for the Price of a Sandwich." *Mozilla Foundation*, February 6, 2024.

Badr, Hanan, and Lena-Maria Möller. "Beyond Mainstream Media and Communication Perspectives on the Arab Uprisings." *Media and Communication* 9, no. 4 (2021): 260–63. https://doi.org/10.17645/mac.v9i4.5151.

Barocas, Solon, Moritz Hardt, and Arvind Narayanan. "Fairness and Machine Learning: Limitations and Opportunities," *fairmlbook.org*, 2019.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜." In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, *FAccT '21*. Association for Computing Machinery, 2021. 610–23. https://doi.org/10.1145/3442188.3445922.

Bengio et al. "Pause Giant AI Experiments: An Open Letter." *Future of Life Institute*, March 22, 2023. https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

Benkler, Yochai, Robert Faris, and Hal Roberts. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press, 2018. https://doi.org/10.1093/oso/9780190923624.001.0001.

Bresnihan, Patrick, and Patrick Brodie. "New Extractive Frontiers in Ireland and the Moebius Strip of Wind/Data." *Environment and Planning E: Nature and Space* 4, no. 4 (2021): 1645–64. https://doi.org/10.1177/2514848620970121.

Brodie, Patrick. "Data Infrastructure Studies on an Unequal Planet." *Big Data & Society* 10, no. 1 (2023). https://doi.org/10.1177/20539517231182402.

Brodie, Patrick, and Julia Velkova. "Cloud Ruins: Ericsson's Vaudreuil-Dorion Data Centre and Infrastructural Abandonment." *Information, Communication & Society* 24, no. 6 (2021): 869–85. https://doi.org/10.1080/1369118X.2021.1909099.

Burkell, Jacquelyn, and Priscilla M. Regan. "Voter Preferences, Voter Manipulation, Voter Analytics: Policy Options for Less Surveillance and More Autonomy." *Internet Policy Review* 8, no. 4 (2019). https://doi.org/10.14763/2019.4.1438.

Carolan, Mary. "RTÉ's Dave Fanning initiates defamation case after his photo appeared in article about sexual misconduct trial of different broadcaster." *The Irish*

Times, January 15, 2024, https://www.irishtimes.com/crime-law/courts/2024/01/15/rtes-dave-fanning-initiates-defamation-case-after-his-photo-appeared-in-article-about-sexual-misconduct-trial-of-different-broadcaster/.

Constantinides, Panos, Ola Henfridsson, and Geoffrey G. Parker. "Introduction—Platforms and Infrastructures in the Digital Age." *Information Systems Research* 29, no. 2 (2018): 381–400. https://doi.org/10.1287/isre.2018.0794.

Couture, Stephane, and Sophie Toupin. "What Does the Notion of 'Sovereignty' Mean When Referring to the Digital?" *New Media & Society* 21, no. 10 (2019): 2305–22. https://doi.org/10.1177/1461444819865984.

Crawford, Kate. *Atlas of AI*. Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press, 2021. https://doi.org/10.12987/9780300252392.

David, Emilia. "Sarah Silverman's lawsuit against OpenAI partially dismissed." *The Verge*, February 13, 2024. https://www.theverge.com/2024/2/13/24072131/sarah-silverman-paul-tremblay-openai-chatgpt-copyright-lawsuit.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." *Preprint, arXiv,* May 24, 2019. https://doi.org/10.48550/arXiv.1810.04805.

Edwards, Benj. "OpenAI says it's "impossible" to create useful AI models without copyrighted material." *Ars Technica*, January 9, 2024, https://arstechnica.com/information-technology/2024/01/openai-says-its-impossible-to-create-useful-ai-models-without-copyrighted-material/.

Estrada, Marcela Suárez, and Sebastián Lehuedé. "Towards a Terrestrial Internet: Re-Imagining Digital Networks from the Ground Up." *Tapuya: Latin American Science, Technology and Society* 5, no. 1 (2022): 2139913. https://doi.org/10.1080/25729861.2022.2139913.

European Commission. "Commission Launches AI Innovation Package to Support Artificial Intelligence Startups and SMEs." *News release*, January 24, 2024, https://digital-strategy.ec.europa.eu/en/news/commission-launches-ai-innovation-package-support-artificial-intelligence-startups-and-smes.

European Commission. "Commission launches calls for contributions on competition in virtual worlds and generative AI." *Press release*, January 9, 2024. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_85.

European Commission. "State of the Union address." September 13, 2023, https://state-of-the-union.ec.europa.eu/state-union-2023_en.

European Commission. “Code of Practice on Disinformation: new reports available in the Transparency Centre.” *Press release*, September 26, 2023. https://digital-strategy.ec.europa.eu/en/news/code-practice-disinformation-new-reports-available-transparency-centre.

Ferrara, Emilio, Herbert Chang, Emily Chen, Goran Muric, and Jaimin Patel. “Characterizing Social Media Manipulation in the 2020 U.S. Presidential Election.” *First Monday*, (October 2020). https://doi.org/10.5210/fm.v25i11.11431.

FTC. “Generative AI Raises Competition Concerns.” Federal Trade Commission, 2023. https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns.

Galperin, Eva and Wafa Ben Hassine. “Changes to Facebook’s "Real Names" Policy Still Don’t Fix the Problem.“ *Electronic Frontier Foundation*, December 18, 2015, https://www.eff.org/deeplinks/2015/12/changes-facebooks-real-names-policy-still-dont-fix-problem.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, et al. “Datasheets for Datasets.” *Preprint, arXiv*, December 1, 2021. http://arxiv.org/abs/1803.09010.

Gingerich, Jonathan. “Is Spotify Bad for Democracy? Artificial Intelligence, Cultural Democracy, and Law.” *Yale Journal of Law and Technology* 24 (2022): 227–316.

Glasius, Marlies, and Marcus Michaelsen. “Authoritarian Practices in the Digital Age| Illiberal and Authoritarian Practices in the Digital Sphere — Prologue.” *International Journal of Communication* 12 (2018):2018. https://ijoc.org/index.php/ijoc/article/view/8899/2459.

Gonçalves, Ana Rita, Diego Costa Pinto, Saleh Shuqair, Marlon Dalmoro, and Anna S. Mattila. “Artificial Intelligence vs. Autonomous Decision-Making in Streaming Platforms: A Mixed-Method Approach.” *International Journal of Information Management* 76 (2024). https://doi.org/10.1016/j.ijinfomgt.2023.102748.

Gray, Mary L., and Suri Siddharth. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Houghton Mifflin Harcourt, 2019.

Gray Widder, David, Sarah West, and Meredith Whittaker. “Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI.” *SSRN Electronic Journal*, 2023. https://doi.org/10.2139/ssrn.4543807.

Grkitsi, Eliza. “Commission ‘compels’ Microsoft to provide information on Bing and generative AI.” *Euractiv*, March 17, 2024. https://www.euractiv.com/section/digital/news/commission-compels-microsoft-to-provide-information-on-bing-and-generative-ai/.

Guess, Andrew M., Neil Malhotra, Jennifer Pan, et al. "How Do Social Media Feed Algorithms Affect Attitudes and Behavior in an Election Campaign?" *Science* 381, no. 6656 (2023): 398–404. https://doi.org/10.1126/science.abp9364.

Guzman, Andrea L, and Seth C Lewis. "Artificial Intelligence and Communication: A Human–Machine Communication Research Agenda." *New Media & Society* 22, no. 1 (2020): 70–86. https://doi.org/10.1177/1461444819858691.

Haimson, Oliver L., and Anna Lauren Hoffmann. "Constructing and Enforcing 'Authentic' Identity Online: Facebook, Real Names, and Non-Normative Identities." *First Monday*, (June 2016). https://doi.org/10.5210/fm.v21i6.6791.

Helming, Clara. "AI Chatbot produces misinformation about elections." *AlgorithmWatch*, December 15, 2023. https://algorithmwatch.org/en/study-microsofts-bing-chat.

Hutchby, Ian. "Technologies, Texts and Affordances." *Sociology* 35, no. 2 (2001): 441–56. https://doi.org/10.1177/S0038038501000219.

Jansen, Bernardus, Natalia Kadenko, Dennis Broeders, Michel Van Eeten, Kevin Borgolte, and Tobias Fiebig. "Pushing Boundaries: An Empirical View on the Digital Sovereignty of Six Governments in the Midst of Geopolitical Tensions." *Government Information Quarterly* 40, no. 4 (2023). https://doi.org/10.1016/j.giq.2023.101862.

Jenka. "AI and the American Smile." *Medium*, March 27, 2023. https://medium.com/@socialcreature/ai-and-the-american-smile-76d23a0fbfaf.

Kabanov, Yury, and Mikhail Karyagin. "Data-Driven Authoritarianism: Non-Democracies and Big Data." In *Digital Transformation and Global Society*. Communications in Computer and Information Science, edited by Daniel A. Alexandrov, Alexander V. Boukhanovsky, Andrei V. Chugunov, Yury Kabanov, and Olessia Koltsova. Springer International Publishing, 2018. https://doi.org/10.1007/978-3-030-02843-5_12.

Kannengießer, Sigrid, and Sebastian Kubitschko. "Acting on Media: Influencing, Shaping and (Re)Configuring the Fabric of Everyday Life." *Media and Communication* 5, no. 3 (2017): 1–4. https://doi.org/10.17645/mac.v5i3.1165.

Kumar, Anilesh, and Daya Thussu. "Media, Digital Sovereignty and Geopolitics: The Case of the TikTok Ban in India." *Media, Culture & Society* 45, no. 8 (2023): 1583–99. https://doi.org/10.1177/01634437231174351.

Lehuedé, Sebastián. "Territories of Data: Ontological Divergences in the Growth of Data Infrastructure." *Tapuya: Latin American Science, Technology and Society* 5, no. 1 (2022). https://doi.org/10.1080/25729861.2022.2035936.

Li, Pengfei, Jianyi Yang, Mohammad A. Islam, and Shaolei Ren. "Making AI Less 'Thirsty': Uncovering and Addressing the Secret Water Footprint of AI Models." *Preprint, arXiv*, April 6, 2023. https://doi.org/10.48550/ARXIV.2304.03271.

Libertson, Frans, Julia Velkova, and Jenny Palm. "Data-Center Infrastructure and Energy Gentrification: Perspectives from Sweden." *Sustainability: Science, Practice and Policy* 17, no. 1 (2021): 152–61. https://doi.org/10.1080/15487733.2021.1901428.

Luccioni, Alexandra Sasha, Sylvain Viguier, and Anne-Laure Ligozat. "Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model." *Preprint, arXiv*, November 3, 2022, https://doi.org/10.48550/arXiv.2211.02001.

Maham, Pegah, and Sabrina Küpert. "Governing General Purpose AI: A Comprehensive Map of Unreliability, Misuse and Systemic Risks." *Policy Brief. Stiftung Neue Verantwortung,* July, 2023. https://www.interface-eu.org/storage/archive/files/snv_governing_general_purpose_ai_pdf.pdf.

Menghani, Gaurav. "Efficient Deep Learning: A Survey on Making Deep Learning Models Smaller, Faster, and Better." *Preprint, arXiv*, June 21, 2021. http://arxiv.org/abs/2106.08962.

Miceli, Milagros, Tianling Yang, Adriana Alvarado Garcia, Julian Posada, Sonja Mei Wang, Marc Pohl, and Alex Hanna. "Documenting Data Production Processes: A Participatory Approach for Data Work." *arXiv.* 11 July, 2022 https://doi.org/10.48550/arXiv.2207.04958.

Micheli, Marina, Marisa Ponti, Max Craglia, and Anna Berti Suman. "Emerging Models of Data Governance in the Age of Datafication." *Big Data & Society* 7, no. 2 (2020). https://doi.org/10.1177/2053951720948087.

Mitchell, Shira, Eric Potash, Solon Barocas, Alexander D'Amour, and Kristian Lum. "Algorithmic Fairness: Choices, Assumptions, and Definitions." *Annual Review of Statistics and Its Application* 8, no. 1 (March 2021): 141–63. https://doi.org/10.1146/annurev-statistics-042720-125902.

Mollen, Anne, Sigrid Kannengießer, Julia Velkova, and Fieke Jansen. *AI Infrastructures and Sustainability: Expanding Perspectives on Automation, Communication and Media*. Palgrave Macmillan, 2025.

Mollen, Anne and Sigrid Kannengießer. "Shaping AI (more) sustainably. Socio-technological perspectives on AI infrastructures." In *Uncertain journeys into digital futures,* edited by Thomas Kox, André Ullrich, and Herbert Zech. Nomos, 2025.

Musiani, Francesca. "*Infrastructuring* Digital Sovereignty: A Research Agenda for an Infrastructure-Based Sociology of Digital Self-Determination Practices." *Information,*

Communication & Society 25, no. 6 (2022): 785–800. https://doi.org/10.1080/1369118X.2022.2049850.

Mytton, David. "Data Centre Water Consumption." *NPJ Clean Water* 4, no. 1 (2021): 11. https://doi.org/10.1038/s41545-021-00101-w.

Nissenbaum, Helen. "A Contextual Approach to Privacy Online." *Daedalus* 140, no. 4 (2011): 32–48. https://doi.org/10.1162/DAED_a_00113.

Novelli, Claudio., Mariarosaria Taddeo, and Luccioni Floridi. "Accountability in artificial intelligence: what it is and how it works. *AI & Society* 39 (2024): 1871–82. https://doi.org/10.1007/s00146-023-01635-y.

noyb. "(Preliminary) noyb WIN: Meta stops AI plans in the EU." *Press release, noyb*, 14 June, 2024, https://noyb.eu/en/preliminary-noyb-win-meta-stops-ai-plans-eu.

O'Donnell, James. "Training AI music models is about to get very expensive." *The Verge*, June 27, 2024, https://www.technologyreview.com/2024/06/27/1094379/ai-music-suno-udio-lawsuit-record-labels-youtube-licensing/?truid=&utm_source=the_algorithm&utm_medium=email&utm_campaign=the_algorithm.unpaid.engagement&utm_content=07-08-2024&mc_cid=982c1b8019.

OpenAI. "How OpenAI is approaching 2024 worldwide elections." *OpenAI blog*, January 15, 2024, https://openai.com/index/how-openai-is-approaching-2024-worldwide-elections/.

Papakyriakopoulos, Orestis, Anna Seo Gyeong Choi, William Thong, et al. "Augmented Datasheets for Speech Datasets and Ethical Decision-Making." In *2023 ACM Conference on Fairness, Accountability, and Transparency*, 881–904. ACM, 2023. https://doi.org/10.1145/3593013.3594049.

Parks, Lisa, and Nicole Starosielski, eds. *Signal Traffic: Critical Studies of Media Infrastructures*. University of Illinois Press, 2015.

Parks, Lisa, Julia Velkova, and Sander de Ridder. *Media Backends Digital Infrastructures and Sociotechnical Relations*. University of Illinois Press, 2023.

Perrigo, Billy. "OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic." *Time Magazine*, January 18, 2023. https://time.com/6247678/openai-chatgpt-kenya-workers/.

Petre, Caitlin, Brooke Erin Duffy, and Emily Hund. "'Gaming the System': Platform Paternalism and the Politics of Algorithmic Visibility." *Social Media + Society* 5, no. 4 (2019). https://doi.org/10.1177/2056305119879995.

Pierri, Paola, and Elizabeth Calderon Lüning. “A Democratic Approach to Digital Rights: Comparing Perspectives on Digital Sovereignty on the City Level.” *International Journal of Communication* (2023): 3600–18.

Pipek, Volkmar, and Volker Wulf. “Infrastructuring: Toward an Integrated Perspective on the Design and Use of Information Technology.” *Journal of the Association for Information Systems* 10, no. 5 (2009): 306–32.

Plantin, Jean-Christophe, and Aswin Punathambekar. “Digital Media Infrastructures: Pipes, Platforms, and Politics.” *Media, Culture & Society* 41, no. 2 (2019): 163–74. https://doi.org/10.1177/0163443718818376.

Png, Marie-Therese. “At the Tensions of South and North: Critical Roles of Global South Stakeholders in AI Governance,” 1434–45. *FAccT '22: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, 2022. https://doi.org/10.1145/3531146.3533200.

Pohle, Julia, and Thorsten Thiel. “Digital Sovereignty.” *Internet Policy Review* 9, no. 4 (2020). https://doi.org/10.14763/2020.4.1532.

Rani, Uma et al. “Development of AI: Role of ‚invisible workers‘ in the AI value chains.” cience-*Policy Brief for the Multistakeholder Forum on Science, Technology and Innovation for the SDGs*, May 2024. https://sdgs.un.org/sites/default/files/2024-05/Rani%3B%20Gobel%3B%20Dhir_Development%20of%20AI.pdf.

Sætra, Henrik Skaug, and John Danaher. “Resolving the Battle of Short- vs. Long-Term AI Risks.” *AI and Ethics*, September 4, 2023. https://doi.org/10.1007/s43681-023-00336-y.

Sahebi, Siavosh, and Paul Formosa. “Social Media and Its Negative Impacts on Autonomy.” *Philosophy & Technology* 35, no. 3 (2022). https://doi.org/10.1007/s13347-022-00567-7.

Sampedro, Víctor, F Javier López-Ferrández, and Patricia Hidalgo. “Digital Disintermediation, Technical and National Sovereignty: The Internet Shutdown of Catalonia’s ‘Independence Referendum.’” *European Journal of Communication* 37, no. 2 (2022): 127–44. https://doi.org/10.1177/02673231211012143.

Savolainen, Laura, and Minna Ruckenstein. “Dimensions of Autonomy in Human–Algorithm Relations.” *New Media & Society* 26, no. 6 (2024): 3472–90. https://doi.org/10.1177/14614448221100802.

Schaul, Kevin et al. “Inside the secret list of websites that make AI like ChatGPT sound smart.” *Washington Post*, April 19, 2023, https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning/.

Schelter, Sebastian, Mozhdeh Ariannezhad, and Maarten De Rijke. "Forget Me Now: Fast and Exact Unlearning in Neighborhood-Based Recommendation." In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2011–15. ACM, 2023. https://doi.org/10.1145/3539618.3591989.

Sedlmeir, Johannes, Alexander Rieger, Tamara Roth, and Gilbert Fridgen. "Battling Disinformation with Cryptography." *Nature Machine Intelligence* 5, no. 10 (2023): 1056–57. https://doi.org/10.1038/s42256-023-00733-2.

Sevignani, Sebastian. *Privacy and Capitalism in the Age of Social Media*. Routledge, 2015. https://doi.org/10.4324/9781315674841.

Sinpeng, Aim. "Digital Media, Political Authoritarianism, and Internet Controls in Southeast Asia." *Media, Culture & Society* 42, no. 1 (2020): 25–39. https://doi.org/10.1177/0163443719884052.

Solaiman, Irene, Zeerak Talat, William Agnew, et al. "Evaluating the Social Impact of Generative AI Systems in Systems and Society." *Preprint, arXiv*, June 12, 2023. http://arxiv.org/abs/2306.05949.

Strubell, Emma, Ananya Ganesh, and Andrew McCallum. "Energy and Policy Considerations for Deep Learning in NLP." *Preprint, arXiv,* 2019. https://arxiv.org/abs/1906.02243.

Surman, Mark et al. "Accelerating Progress Toward Trustworthy AI." *Mozilla Foundation*, 2024. https://foundation.mozilla.org/en/research/library/accelerating-progress-toward-trustworthy-ai/whitepaper/.

Susser, Daniel, Beate Roessler, and Helen Nissenbaum. "Technology, Autonomy, and Manipulation." *Internet Policy Review* 8, no. 2 (2019). https://doi.org/10.14763/2019.2.1410.

Tar, Julia. "EU Commission to examine Microsoft-OpenAI partnership." *Euractiv*, January 9, 2024. https://www.euractiv.com/section/competition/news/eu-commission-to-examine-microsoft-openai-partnership/.

Tar, Julia, "Apple's ChatGPT integration raises data privacy and competition questions." *Euractiv*, June 17. 2024, https://www.euractiv.com/section/data-privacy/news/apples-chatgpt-integration-raises-data-privacy-and-competition-questions/.

The New York Times. "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work." *New York Times*, December 27, 2023. https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.

Trepte, Sabine. "The Social Media Privacy Model: Privacy and Communication in the Light of Social Media Affordances." *Communication Theory* 31, no. 4 (2021): 549–70. https://doi.org/10.1093/ct/qtz035.

Van Dijck, José. "'You Have One Identity': Performing the Self on Facebook and LinkedIn." *Media, Culture & Society* 35, no. 2 (2013): 199–215. https://doi.org/10.1177/0163443712468605.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, et al. "Attention Is All You Need." *Preprint, arXiv*, June 2, 2017. https://doi.org/10.48550/ARXIV.1706.03762.

Vincent, James. "OpenAI co-founder on company's past approach to openly sharing research: 'We were wrong'." *The Verge*, March 15, 2023, https://www.theverge.com/2023/3/15/23640180/openai-gpt-4-launch-closed-research-ilya-sutskever-interview.

Vipra, Jai, and Anton Korinek. "Market Concentration Implications of Foundation Models: The Invisible Hand of ChatGPT." *Center on Regulation and Markets Working Paper #9. Brookings*, 2023. https://www.brookings.edu/wp-content/uploads/2023/09/market-concentration-implications-of-foundation-models-final-1.pdf.

Zuboff, Shoshana. "Surveillance Capitalism and the Challenge of Collective Action." *New Labor Forum* 28, no. 1 (January 2019): 10–29. https://doi.org/10.1177/1095796018819461.