

Secondary Publication



Jungherr, Andreas

Artificial Intelligence and Democracy : A Conceptual Framework

Date of secondary publication: 11.08.2023

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-899890

Primary publication

Jungherr, Andreas: Artificial Intelligence and Democracy : A Conceptual Framework. In: Social Media + Society 9 (2023), 3, pp. 1-14. DOI: 10.1177/20563051231186353

Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holder(s).

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by-nc/4.0/legalcode>

Artificial Intelligence and Democracy: A Conceptual Framework

Social Media + Society
July-September 2023: 1–14
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20563051231186353
journals.sagepub.com/home/sms



Andreas Jungherr

Abstract

The success and widespread deployment of artificial intelligence (AI) have raised awareness of the technology's economic, social, and political consequences. Each new step in the development and application of AI is accompanied by speculations about a supposedly imminent but largely fictional artificial general intelligence (AGI) with (super-)human capacities, as seen in the unfolding discourse about capabilities and impact of large language models (LLMs) in the wake of ChatGPT. These far-reaching expectations lead to a discussion on the societal and political impact of AI that is largely dominated by unfocused fears and enthusiasms. In contrast, this article provides a framework for a more focused and productive analysis and discussion of AI's likely impact on one specific social field: democracy. First, it is necessary to be clear about the workings of AI. This means differentiating between what is at present a largely imaginary AGI and narrow artificial intelligence focused on solving specific tasks. This distinction allows for a critical discussion of how AI affects different aspects of democracy, including its effects on the conditions of self-rule and people's opportunities to exercise it, equality, the institution of elections, and competition between democratic and autocratic systems of government. This article shows that the consequences of today's AI are more specific for democracy than broad speculation about AGI capabilities implies. Focusing on these specific aspects will account for actual threats and opportunities and thus allow for better monitoring of AI's impact on democracy in an interdisciplinary effort by computer and social scientists.

Keywords

artificial intelligence, democracy, self-rule, equality, elections, autocracy

Artificial Intelligence and Democracy

The success and widespread deployment of artificial intelligence (AI) have raised awareness of the technology's economic, social, and political consequences. The most recent step in AI development—the application of large language models (LLMs) and other transformer models to the generation of text, image, video, or audio content—has come to dominate the public imaginary of AI and accelerated this discussion. But to assess AI's societal impact meaningfully, we need to look closely at the workings of the underlying technology and identify the areas of contact within fields of interest. This article proposes a corresponding conceptual framework for the study of the current and prospective impact of AI on democracy.

AI has become a pervasive presence in society. Recent technological advances have allowed for the broad deployment of AI-based systems in many different areas of social, economic, and political life. In the process, AI has had, or is expected to have, a strong effect on each area it touches. We see examples in discussions about the algorithmic

shaping of digital communication environments and the associated deterioration of political discourse (Kaye, 2018); the flooding of the public arena with false or misleading information enabled through generative AI (Krebs et al., 2022); algorithmically stimulating political conflict (Settle, 2018); AI's impact on international human rights law (Gellers & Gunkel, 2023); the future of work and AI's role in the replacement of jobs and related automation-driven unemployment (Acemoglu & Johnson, 2023; Brynjolfsson & McAfee, 2016; Frey, 2019); and AI's impact on shifting the competitive balance between autocracies and democracies (Filgueiras, 2022; Lee, 2018). With these developments, AI has also begun to touch the very idea and practice of democracy.

University of Bamberg, Germany

Corresponding Author:

Andreas Jungherr, Chair for Political Science, especially Digital Transformation, Institute for Political Science, University of Bamberg, Feldkirchenstraße 21, 96052 Bamberg, Germany.
Email: andreas.jungherr@gmail.com



Current adaptation of AI resembles prior waves of technological change and their political impact. Technologies provide supporting structures for the coordination of social, economic, and political life. Through their design, underlying mechanisms, and inputs and outputs, different technologies influence the societal fields and processes mediated through them or on which they rely (Winner, 1980). Technology and technological shifts therefore have effects on politics and political competition by asymmetrically favoring actors, factions, or groups depending on their alignment or misalignment with the affordances emerging from the technology of the day (Bimber, 2003; Castells, 2009/2013; Jungherr et al., 2019; Müller, 2021). In the past, we saw this with the impact of soil maps, geometry, and writing (Stasavage, 2020), the printing press (Eisenstein, 1979; Kaufmann, 2019), newspapers (Schudson, 1978), broadcast media (Neuman, 1991; Prior, 2007), and, more recently, digital media (Jungherr et al., 2020; Williams & Carpinini, 2011). The growing presence of AI-based systems in society demands an interrogation, along similar lines, of how AI affects the idea and practice of democracy (Risse, 2023).

AI is often discussed as a threat to society if not human life itself (Bostrom, 2014)—a discussion predicated on a largely imaginary *artificial general intelligence* (AGI) able to autonomously perceive, reason, decide, and act in varying contexts with human or superhuman capabilities. This notion, derived mostly from speculative fiction, has little correspondence with AI-based systems currently deployed or lab research on the development of AI (Agrawal et al., 2018/2022; Larson, 2021; M. Mitchell, 2019; Smith, 2019). In fact, actually existing AI is predominantly *narrow AI* trained on domain-specific data to perform domain-specific tasks (M. Mitchell, 2019, p. 45f.). Accordingly, in examining the impact of AI on democracy, it is important not to get sidetracked by imaginary AGI and instead focus on specific instances of narrow AI, the conditions for its successful deployment, its uses in specific areas of interest, and their effects.

This article provides answers to two questions:

1. What type of AI do we encounter when examining AI's impact on democracy?
2. How and in what respects does this touch on the idea or practice of democracy?

The article provides a novel conceptual framework for assessing and monitoring how AI—even in its current narrow manifestation—affects the idea and practice of democracy. It combines a discussion of AI's constitutive technical features with its expected impact on democracy, informed by political theory and empirical findings from various fields. The broadness of this account will of course lead to a lack of nuance in specific areas. While this might feel unfortunate to specialists, lack of nuance can be a feature of social theory, allowing for the development of abstract frameworks, affording in turn the theory-informed charting of

new subjects and the establishment of connections between fields (Healy, 2017). While social scientists and computer scientists alike may feel shortchanged, the combination of both their perspectives promises to inform both groups. For technologists, the article grounds their discussion of AI's impact on democracy in political theory. For social scientists, it connects their discussion with the actual workings of available AI technology. Some of the examples presented here will remain somewhat speculative. Again, this might be seen like a detriment by some. But if we accept that current technological developments—like AI—will come to shape and impact democratic practice, ideas, and potentially even structures, we must allow ourselves to speculate and work with thought experiments. Simply relying on data and examples documenting the present and the past will render us blind to the evolving nature of democracy and leave us silent on questions of design choices at the interface between technology and democracy (Papacharissi, 2021).

Artificial Intelligence

AI has a rich history of approaches, applications, and associated imaginaries (McCorduck, 2004; Nilsson, 2010; Russell & Norvig, 1995/2021). To develop specific expectations regarding AI's impact on democracy, we must be precise about the term and how the underlying technology works. There is a wide variety of competing definitions, but for the purposes of this article, AI can be defined as “the study and construction of agents that do the right thing” (Russell & Norvig, 1995/2021, p. 22). More specifically, AI has been described as:

“[T]hat activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment.” (Nilsson, 2010, p. xiii)

This includes approaches that allow machines to pursue tasks, sometimes on par with and sometimes surpassing the ability of humans. The idea of a powerful machine-intelligence has inspired far-reaching expectations and fears regarding the potential or threats associated with AI, ranging from economic growth (Brynjolfsson & McAfee, 2016) and post-human transcendence (Hanson, 2016) to a downright menace to human existence (Bostrom, 2014). Expectations of AI's supposed general ability to address tasks or form decisions across multiple domains (AGI) are extrapolated from AI's very real success in data-driven completion of specific tasks in specific domains (*narrow AI*). This, though, is a category error.

AI's current successes do not belong to a still largely fictional AGI but rely instead on advances in narrow AI. The technology has succeeded in completing specific tasks in various domains. These successes depend strongly on advances in the identification of patterns that link available

data points to outcomes of interest. On a very basic level, this can be the application of simple regression or machine-learning models that can be surprisingly successful in many tasks and are often characterized as a form of AI. But, of course, they are far from the understanding of AI introduced above. Truer to that understanding of AI are more advanced approaches such as deep learning (Goodfellow et al., 2016) and reinforcement learning (Sutton & Barto, 2018). These approaches have been successful in many different contexts, including computer vision, machine translation, medical diagnosis, robotics, and voice recognition (LeCun et al., 2015). Examples of their promise include predictions of possible but yet unknown biological or chemical compounds (Chow et al., 2018; King et al., 2009; Schneider, 2018) or strategic action in game play (Bakhtin et al., 2022; Silver et al., 2016, 2017, 2018). Recently, transformer models, a subclass of deep learning, have proven to be highly promising (Parmar et al., 2018; Vaswani et al., 2017) in the autonomous generation of text, image, and video content (Brown et al., 2020; Ramesh et al., 2022). As we will see below, these advances are starting to feature in politics and might come to impact democracies.

But these approaches do not work universally. In fact, they depend on a set of preconditions, which might limit their uses in politics. Some are obvious. For example, to be successful AI needs to be able to access some digital representation of its environment, either through sensors mapping the world or through the input of existing data. Where these representations are difficult to come by or data are scarce, as in many areas of politics, AI will not be successful. Other preconditions are not so obvious. For example, for AI to produce helpful results, the underlying connections between inputs and outputs must be stable over time. This points to two problems: unobserved temporal shifts between variables (Lazer et al., 2014) and the dangers of relying on purely correlative evidence without the support of causal models (Pearl, 2019; Schölkopf et al., 2021).

More important still, especially with respect to democracy, is that normatively speaking the past must provide a useful template for the future. Change is a crucial feature of societies, especially the extension of rights and the participation of previously excluded groups. Over time, many societies strive to decrease discrimination and increase equality. In fact, many policies are consciously designed to break with past patterns of discrimination. AI-based predictions and classifications based on past patterns risk replicating systemic inequalities and even structural discrimination (Bolukbasi et al., 2016; Christian, 2020; S. Mitchell et al., 2021).

Problems that share these characteristics can be found in many areas, such as the digital economy, commerce, digitally mediated social interactions, robotics, and sports. AI has proven highly successful in these areas. At the same time, few problems in politics and democracy more broadly share these characteristics. This limits the application of AI in society and, accordingly, its impact on democracy.

The Impact of Artificial Intelligence on Democracy

AI's recent successes and its broad deployment in many areas of social, economic, and political life have begun to raise questions regarding whether and how AI impacts democracy. The idea and practice of democracy are highly contested concepts with competing accounts of great nuance. The associated discussions within political theory are highly productive and successful in identifying different normative, procedural, or structural features and consequences within our understanding of democracy (Dahl, 1998; Guttman, 2007; Landemore, 2012; Przeworski, 2018; Tilly, 2007). Still, for the purposes of this article, it is necessary to reduce this rich field to a few important—if sometimes contested—features of democracy in which we can expect AI impact. What the argument might lose in nuance, is compensated by the provision of a broadly applicable conceptual framework.

This article presents four areas of impact at different analytical levels: at the *individual* level, AI impacts the conditions of self-rule and people's opportunities to exercise it; at the *group* level, AI impacts equality of rights among different groups of people in society; at the *institutional* level, AI impacts the perception of elections as a fair and open mechanism for channeling and managing political conflict; and at the *systems* level, AI impacts competition between democratic and autocratic systems of government.

Level	Area of impact
Individual	Self-rule
Group	Equality
Institutional	Elections
System	Competition between systems

This article does not follow one specific theory of democracy, instead approaching democracy as a multifaceted phenomenon (Asenbaum, 2022). This offers a broader view of the areas where AI might conceivably affect society in ways relevant to the performance and quality of democracy.

Artificial Intelligence and Self-rule

One tenet of democracy is that governments should be chosen by those they will serve. Such self-rule is a normative idea about legitimizing the temporal power of rulers over the ruled and a practical idea that distributed decision-making is superior to other more centralized forms of decision-making or rule by experts (Dahl, 1998; Landemore, 2012; Landemore & Elster, 2012; Schwartzberg, 2015). AI impacts both the ability of people to achieve self-rule and the perceived superiority of distributed decision-making over expert rule in complex social systems, highlighting potential limits to self-rule in several ways.

Shaping Information Environments. The legitimacy of self-rule is closely connected with the idea of people being able to make informed decisions for themselves and their communities. This depends at least in part on the information environment in which they are embedded (Jungherr & Schroeder, 2022). AI affects these informational foundations of self-rule directly. This includes how people are exposed to and can access political information, can voice their views and concerns, and how these informational foundations potentially increase opportunities for manipulation (Jungherr & Schroeder, 2023).

Algorithmic shaping of digital information environments based on people's inferred information preferences or predicted behavioral responses (Narayanan, 2023) has raised particularly strong concerns (Kaye, 2018). Key among these is that people will be exposed only to information with which they are likely to agree, thus losing sight of the other political side. Empirical findings suggest that these fears may be overblown (Flaxman et al., 2016; Kitchens et al., 2020; Scharkow et al., 2020). In fact, in digital communication environments, people may encounter more political information about the other side and that they disagree with than in other information environments. This can be a problem, especially for political partisans, because it increases the salience of political conflict (Settle, 2018). But the degree to which this mechanism is driven by AI or might even be lessened through specific algorithm design remains as of now unknown.

Going further, several authors have diagnosed various ill effects of digital communication environments on information quality and political discourse, some AI-driven and others independent of AI (Bennett & Livingston, 2021). While clearly important, these diagnoses risk overestimating the quality of prior information environments and the role of information for people in their exercise of self-rule. In fact, critiques of the quality of media in democracies abounded well before digital media became prevalent (Keane, 2013).

In addition, most people do not follow the news closely, do not hold strong political attitudes, and do not perform well when tested on their political knowledge (Converse, 1964; Lupia & McCubbins, 1998; Prior, 2007; Zaller, 1992). They seem to rely on informational shortcuts or on social structures to exercise self-rule (Achen & Bartels, 2016; Kuklinski & Quirk, 2000; Lodge & Taber, 2013; Popkin, 1991). Hence, these mechanisms can also be expected to mediate the impact of AI-driven shaping of information environments. To assess AI's impact fully, research needs to consider not only information environments but must also look at whether and how AI affects the structural and social factors that mediate the impact of political information on self-rule.

It does not appear that AI-driven shaping of digital information environments inevitably leads to a deterioration of access to information necessary for people to exercise their right to self-rule. Nevertheless, there is much opaqueness in

the way digital communication environments are shaped. The greater the role of these environments in democracies, the greater the need for assessability of the role of AI in their shaping (Jungherr & Schroeder, 2023). We also need regular external audits of the effects of AI on the information visible on online platforms, especially the nature and kind of information that is algorithmically promoted or muted.

Economics of News. AI might also come to indirectly impact the creation and provision of relevant political information by changing the economic conditions of news production. For one, recent successes in the development of transformer models suggest that AI might soon be used by media providers to automatically generate text, image, or video content. This might lead to an acceleration of existing trends toward automated content generation in news organizations (Diakopoulos, 2019). This puts pressure on journalists who might see routine tasks shift toward AI-enabled systems but also on news organizations who might face a new set of ultra low-cost competitors who specialize on automatically generated news content. This potentially increases pressure on journalists' salaries as well as the audiences and profits of news companies, intensifying existing pressures on news as a business (Nielsen, 2020).

In addition, AI reconfigures the way news and political information are accessed by the public. Search engines like Bing and Google are experimenting with LLMs to provide users with automatically generated content in reaction to search queries instead of links to content provided by news and information providers. This limits monetization opportunities for small- or middle-sized media organizations without strong brand identity and loyalty, which in the past could generate traffic based on query-based referrals from search engines or social networking sites. These new limitations on monetization opportunities might lead to a decline in the coverage of politics, or even a reduction in the number of news organizations. This in turn would limit the total amount and diversity of information available to people to develop informed decisions. This will hit political outsiders and challengers the hardest who rely on smaller information providers for coverage. This decline in monetization opportunities of news will thus likely lead to a strengthening of existing institutions, media brands, and associated power relations (Jungherr & Schroeder, 2023).

In addition, public perceptions of digital communication environments being dominated by AI-generated content—some of it correct, some of it actively misleading, some of it accidentally misleading—might contribute among parts of the population to an increased valuation of select news organizations, whose process of news production and quality insurance they have come to trust. These news brands might thus find themselves strengthened through an increase of AI-generated content in open communication environments or in the coverage by cost-cutting competitors. Of course,

this expectation only holds if these news brands are seen as providing added value over AI-generated content.

It is also important to remember that this AI-driven turn to specific news brands is only likely to hold for audience members who engage with news and politics demanding accurate information and those interested in politics. This will likely be socio-economically well-resourced and politically engaged people (Prior, 2018; Schlozman et al., 2018). Others might feel fine with free or automatically generated content. This is likely to reinforce an informational divide between politically interested and disinterested audiences that already has grown following the switch from a low-choice mass media environment to high-choice digital communication environments (Prior, 2017). In countries without strong public broadcasters, like the US, this divide will also run along economic lines, allowing those able to pay for news to access high-quality, curated, and quality-checked information, while leaving those not able (or willing) to pay to the noisy, (partially) automated, and contested free digital information environment. Over time, this might mean that socio-economic divides decide (or are seen to decide) over the ability of people to come to informed political decisions.

Speech. AI does not only impact access to information, it also affects the expression of opinions, interests, and concerns in digital communication environments. With digital communication environments becoming increasingly areas for the expression of voice, surfacing of concerns, and construction of political identities, this is an important element in AI's shaping of the conditions for self-rule.

The perceived ability of AI to classify content has put it at the forefront of the fight against harmful digital speech and misinformation. AI is used broadly by tech companies to classify user content to stop it from publication or flag it for moderation (Douek, 2021; Kaye, 2018). Details of the applied procedures, their successes, and error rates are opaque to outsiders, making it difficult to assess the broadness of AI's uses and its effects on speech. This is problematic: harmful speech and misinformation are both difficult categories for classification. Neither category is objective nor stable and both require interpretation as meaning shifts across contexts and time. This makes them difficult to identify with automated data-driven AI and risks suppression of legitimate political speech.

In addition, the technical workings of AI also impact the type of speech becoming visible in AI-shaped spaces. By learning typical patterns within a given set of cases, AI will lean toward averages. For AI-enabled shaping and summarizing of speech or political positions, this will favor common positions, concerns, and expressions. Outsiders and minority positions, concerns, and expressions will in unadjusted AI-shaped communication environments end up submerged and become invisible. AI would thus negatively impact the ability of a society to make itself visible to itself,

lower democracies' information processing capacities, and strengthen the political status quo (Jungherr & Schroeder, 2023).

Still, there are few alternatives to AI-based moderation given the pure volume of content being published in digital communication environments (Douek, 2021), which makes it important to gain a better understanding of AI-based moderation's workings and effects. Accordingly, AI-based moderation needs assessability provided by platforms and external audits to ensure its proper workings.

AI-based moderation, however, is not only a risk. Scholars and commentators have long pointed to the limits of large-scale political deliberation imposed through inefficiencies in information distribution, surfacing of preferences, and coordination of people. AI may improve on some of these inefficiencies by predicting individual preferences, classifying information, and shaping information flows (Landmore, 2022). This in turn might open opportunities for new deliberative and participatory formats in democracies, thereby strengthening and vitalizing democracy.

It is important to remain aware of both the risks and the opportunity AI provides for moderating speech and surfacing concerns in digital communication environments. AI can contribute to creative solutions to some of the technical challenges underlying successful self-rule. But if it is to do so, we need to know more about its actual uses, effects, and risks. This demands greater transparency from digital platforms and continued vigilance and attention from civil society.

Manipulation. AI could also negatively impact individual informational autonomy by predicting the reactions of people to communicative interventions. This could allow professional communicators to reach people in exactly the right way to shift opinions and behavior. Sanders and Schneier (2021) present a thought experiment that illustrates how lobbyists might use AI to predict the likelihood of success of bills they introduce to legislators. While still far from realization, their example shows interested parties employing AI to increase the resources available to them and potentially targeting interventions aimed at influencing people to behave in ways beneficial to those same parties. AI can also be used to generate messages aimed at persuading people, with early working papers indicating interventions designed by LLMs to have persuasive appeal (Bai et al., 2023). Similarly, LLMs are currently used by academics and campaign professionals to simulate reactions and attitudes by prototypical voters for message testing and research, although the precision and validity of these approaches are contested (Bisbee et al., 2023; Horton, 2023; Kim & Lee, 2023).

Fears also exist regarding people encountering targeted communicative interventions in digital communication environments. By predicting how people might react to an advertisement, digital consultancies could use AI to tailor interventions to influence people. The English consultancy firm Cambridge Analytica, which claimed to be able to

predict which piece of information displayed on Facebook was necessary to get people to behave in ways beneficial to its electoral clients, provided a first taste of this problem. While the company's claims have been debunked (Jungherr et al., 2020, pp. 124–130), the episode speaks to the perception of AI's power to manipulate people at will, as well as the willingness of journalists and the public to accept widely exaggerated claims about the power of digitally enabled manipulation irrespective of contradicting evidence.

Recent advances in transformer models have opened new avenues for potential manipulation through the automated production of text or images (Brown et al., 2020; Ramesh et al., 2022). There are legitimate uses of these models, as well as nefarious ones. For instance, they facilitate the automated generation of content based on raw information or event data, as found in sports coverage or the stock market (Diakopoulos, 2019). This is largely unproblematic since AI translates information from one form of representation—such as numerical or event data—into another—such as a narrative news article.

More problematic are cases in which AI does not simply translate one representation of information into another but generates content based on prompts and past patterns. Examples include text or image responses to textual prompts in the form of questions or instructions. AI has no commitment to the truth of an argument or observation; it is only imitating their likeness as found in past data. Today's AI is committed only to the representation of the world, an object, or an argument available to it, not to the world, object, or argument as such (Smith, 2019). Thus, AI output taken at face value cannot be trusted because it is not necessarily true, only plausible.

More problematic still is the chance that future AI could be used to produce fake information at scale. This could take the form of targeted fakes aimed at misleading people, or flooding information environments with masses of unreliable or misleading AI-generated content. This would dilute information environments, making it more difficult for people to access crucial information and/or making information appear untrustworthy.

At the same time, somewhat counterintuitively, a mass-seeding of automated misinformation might also contribute to the strengthening of professional news and information curation discussed above. When the prevalence of unreliable or misleading information in digital communication environments becomes evident, the premium for reliable information rises. Accordingly, professional, reliable, and impartial news sources might see a reversal of fortune compared to their economic and ideational challenges of the last 20 years. This way, automated misinformation in scale might turn out to strengthen intermediary institutions that provide information in democracies.

It is important to note that these uses of AI are still projected and may not come to pass given the limits of the underlying technology, the development of efficient countermeasures, and/or the persistence of mediating structures that

limit the effects of information overall. But considering recent technological advances, these uses have come to feature strongly in the public imagination and demand for critical reflection by social and computer scientists.

Expert Rule. Support for self-rule is also closely connected with the assessment of expert rule being limited in complex social systems. Expertise is important, but has limited predictive power in complex societies, and the decentralized decision-making and preference surfacing of self-rule, while imperfect, are seen as superior for settling on collectively binding decisions (Dahl, 1998; Lindblom, 2001). The growing availability of data in ever more domains, coupled with new analytical opportunities offered by AI, have raised hopes for new predictive capabilities in complex societies (Kitchin, 2014). AI not only highlights the weaknesses of people making political decisions but also increases the power of experts.

AI brings new opportunities in the modeling and prediction of societal, economic, ecological, and geopolitical trends, promising to provide experts with predictions of people's behavior in reaction to regulatory or governance interventions. While the actual quality of these approaches is still open to question, they have strong rhetorical and legitimizing power. They increase the power of experts, who—sometimes actually and sometimes rhetorically—rely on AI-supported models to ground their advice on how societies should act considering major societal challenges. This apparent increase in the power of experts to guide societies in responding to challenges can reduce the option space available for democratic decision-making, shifting the question from whether people *can* to whether they *should* decide for themselves. In this, AI could induce a transition from self-rule to expert rule and thereby weaken democracy.

Power of Technology Companies. AI also increases the power of firms over the public and even over states. While the theoretical breakthroughs in the current wave of AI began at universities, it is firms that lead in their practical application, further development, and broad rollout (Ahmed et al., 2023; Metz, 2021). Over time, the power to innovate and critically interrogate AI may shift from public to commercial actors, weakening AI oversight and regulation by democratically legitimated institutions. These challenges can be clearly seen in attempts by both the US and EU at getting to grips with regulating AI development and uses (Criddle & Murphy, 2023; Espinoza & Johnston, 2023).

There is also the issue of economic and political power. AI has allowed companies such as Google and Amazon to dominate multiple economic sectors (Bessen, 2022; Brynjolfsson et al., 2023). Governments have also begun to rely on AI-based service providers to support executive functions such as policing and security. The result is a growing government dependence on AI companies and an opaque transfer of knowledge from governments to these service

providers. Add to this power over AI-enabled information flows and governance over political speech (Jungherr & Schroeder, 2023), and AI companies hold central positions in democracies, potentially negatively influencing the abilities of people for self-rule. This shows the importance of effective government and civil society oversight of companies that provide AI and those that employ AI to ensure that the foundations of meaningful self-rule hold as societies begin to rely more on AI-supported systems.

Artificial Intelligence and Equality

Democracy depends on people having equal rights to participation and representation (Dahl, 1998). While this ideal is imperfectly realized and strongly contested in practice (Phillips, 2021; Young, 2002), democracies are in an ongoing struggle to extend rights to formerly excluded groups. AI's reliance on data documenting the past risks subverting this process and instead continuing past discrimination into the future, thereby weakening democracy.

By predicting how people will behave under various circumstances based on observations from the past, AI differentiates among people based on criteria represented in data points. This risks reinforcing existing biases in society and even porting socially, legally, and politically discontinued discriminatory patterns into the present and future (Eubanks, 2018; Mayson, 2019; Mehrabi et al., 2022; S. Mitchell et al., 2021; Obermeyer et al., 2019). This makes continuous observation and auditing of AI implementation crucial.

People's visibility to AI depends on their past representation in data. AI has trouble recognizing those who belong to groups underrepresented in the data used to train it. For example, minorities not traditionally represented in data sets will remain invisible to computer vision (Buolamwini & Gebru, 2018), and historically underrepresented groups will not be associated with specific jobs and thereby risk discrimination in AI-assisted job procedures (Caliskan et al., 2017). This general pattern is highly relevant to democracy: for example, the systematic invisibility of specific groups means they would be diminished in any AI-based representation of the body politic and in predictions about its behavior, interests, attitudes, and grievances. Accordingly, already disenfranchised people could risk further disenfranchisement and discrimination in the roll out of government services, the development of policy agendas based on digitally mediated preferences and voice, or face heightened persecution from the state security apparatus.

AI also makes some people *more* visible. Historically, marginalized groups will be overrepresented in crime records, negatively impacting group members in AI-based approaches to policing or sentencing (Chouldechova, 2017; Christian, 2020; Ferguson, 2017). In countries like the US, where voting rights are withheld for felons to varying degrees depending on state jurisdiction, systematic biases in AI-supported policing and sentencing might over time come to systematically

bias the electorate against historically disenfranchised groups (Aviram et al., 2017). AI-based approaches can also have a profound effect on electoral redistricting (Cho & Cain, 2020). AI could lead to a reinforcement of structural inequality and discrimination by continuing patterns found in historical data even if a society is trying to enact more equal, less discriminatory practices.

Extrapolating from this, we can expect subsequent AI-based representations of public opinion, the body politic, and AI-assisted redistricting to be biased against groups marginalized in the past. Different degrees of visibility to AI could increase the democratic influence of some groups and decrease that of others. For instance, AI might contribute to an increase of resources for the already privileged by making their voices, interests, attitudes, concerns, and grievances more visible and accessible to decision-makers. AI might use the preferences of visible groups in predictions about political trends and policy impact while ignoring those of less visible groups.

AI can also have adverse effects on the labor market. While in principle firms could invest in automation to allow workers to pursue new tasks and thereby increase the value of their labor, it appears that firms do so mostly to lower their own labor costs by substituting AI for human labor-based tasks (Acemoglu & Restrepo, 2019). This lowers workers' bargaining power and income by substituting labor for capital, which in turn threatens to increase economic inequality and weaken workers' collective bargaining power. Consequently, this could also lower workers' political influence and representation (Acemoglu, 2021; Gallego & Kurer, 2022).

What type of labor is affected by AI-based technological progress, although, is uncertain. Automation traditionally substitutes for routine human tasks and thus affects mostly low-skilled workers (Acemoglu & Restrepo, 2022b; Frey, 2019). But subsequent waves of AI innovation have shown that routine tasks underly many professions, including white-collar and knowledge ones long perceived as being immune to automation. The impact of AI in changing the political fortunes of workers might thus concern larger groups in the economy than traditional forms of automation. This can already be seen in the current discussion about the impact of LLMs and generative AI on the creative and software industry, which until now seemed to be exempt from the dangers of automation-driven job replacement. These emerging fault-lines can already be seen in the Hollywood writers' strike from 2023, in which screenwriters demanded contractual protection against studio uses of AI for writing tasks (Wilkinson, 2023).

At the same time, AI can help aging societies complete substitutable work tasks and concentrate the shrinking labor force on currently nonsubstitutable tasks, thereby maintaining productivity levels in the face of growing demographic pressures in several developed economies (Acemoglu & Restrepo, 2022a). But realizing AI's economic potential for societies

means ensuring that respective gains are broadly shared and do not only benefit a narrow elite. Especially with prosperity gains from digital technology, this link of shared prosperity gains seems to be broken. This raises concerns as to whether elites manage to capture respective AI-enabled gains while most people only face automation-driven economic risks (Acemoglu & Johnson, 2023). This would increase inequality in society and weaken democracy. This potentially dangerous development puts the specifics of AI's implementation and its public and regulatory oversight into focus.

AI clearly touches on equality within democracies. Inequalities might arise in the allocation of options and state services using AI-based systems, people's visibility and representation within AI-based systems, and the provision or withdrawal of economic opportunities for people whose job tasks can be replaced with AI. These are, therefore, important areas for further interrogation and, if necessary, regulatory intervention.

Artificial Intelligence and Elections

Democracies rely on elections, which channel and manage political conflict by providing factions the opportunity to gain power within an institutional framework. This works only if each faction sees a genuine opportunity to win power (Przeworski, 2018), making democracy a system of "organized uncertainty" (Przeworski, 1991, p. 13). AI applications threaten to offset this perceived uncertainty of who will lose and who will win elections. However, the uses of AI in this field are limited.

Data-driven approaches are limited in the prediction of individual voters' behavior. While the voting behavior of committed partisans can be predicted with some probability (Hersh, 2015; Nickerson & Rogers, 2014)—at least in two-party systems—predicting the behavior of people who are only weakly involved with politics is much harder. People do not always vote, and when they do the context can vary greatly. Their vote choices are for the most part not available to modelers, making predicting voting behavior automatically a problem for which AI is not well suited. The uncertainty of election victories will thus remain for the foreseeable future. But campaigns can develop other relevant data-driven models of elections, such as someone's probability of voting or donating money (Hersh, 2015; Issenberg, 2012; Nickerson & Rogers, 2014), which could give campaigns a competitive advantage. Any such advantage is likely fleeting, though, given the broad availability of AI-based tools and campaign organizations learning from others' successes and failures (Kreiss, 2016).

Firms and governments might also seek to use AI to predict election outcomes or the electorate's mood swings and possibly intervene. These efforts are limited by the same challenges raised above, but the public impression of this capability might be enough to undermine and delegitimize elections and give election losers a pretext to challenge results rather than conceding.

Cambridge Analytica's supposed role in the United Kingdom's Brexit vote and the 2016 US presidential election previewed some of the challenges. While there is little indication that data-based psychological targeting was widely used or had sizable effects, these episodes still loom large in the public imagination as an example of AI's perceived power in election manipulation (Jungherr et al., 2020, pp. 124–130). We can expect widespread AI use in economic, political, and social life to shift people's expectations of its uses and abuses in electioneering, irrespective of its actual uses or inherent limitations.

Overall, AI's impact on elections seems limited, given the relative scarcity of the predicted activity—voting. While indirect effects are possible through potential opportunities for competitive differentiation, it is doubtful that this can translate into a consistent, systemic shift of power, given the broad availability of AI tools. More likely is the indirect impact mentioned above: that by transposing expectations regarding AI's supposed powers from industry and science to politics, the public may come to believe that AI is actually able to offset the "organized uncertainty" of democratic elections. This alone could weaken public trust in elections and acceptance of election results. It is thus important to keep organized uncertainty alive in the face of AI, not weaken it through irresponsible and fantastical speculation.

Artificial Intelligence and the Autocratic Competition

AI also affects the relationship between democracy and other systems of governance, such as autocracy, which some have argued has an advantage in the development and deployment of AI. Firms and governments that in democracies face limits to AI deployment or pervasive data collection about people's behavior have more leeway in autocracies. A close connection between the state and firms developing and deploying AI in autocracies creates an environment of permissive privacy regulation that provides developers and modelers with vast troves of data, allowing them to refine AI-enabled models of human behavior. Add centrally allocated resources and training of large numbers of AI-savvy engineers and managers, and some expect the result to be a considerable competitive advantage in developing, deploying, and profiting from AI-supported systems (Figueiras, 2022; Lee, 2018). This may allow for asymmetric developmental progress in AI, state capacity, economic benefits, and potentially even military prowess favoring autocracies over democracies.

Leaving aside normative considerations, democracies have been seen, on a purely functional level, to be superior to autocracies due to their superior performance as information aggregators and processors (Kuran, 1995; Lindblom, 1965; Ober, 2008; Wintrobe, 1998). Free expression, a free press, and electorally channeled competition between factions provide democracies with structural mechanisms that surface information about society, the actions of bureaucracies, and the impact of policies. In contrast, autocracies restrict information flows by controlling speech, the media,

and political competition, leaving governments in the dark regarding local situations, the preferences of the public, the behavior or corruption in their bureaucracies, and ultimately, the consequences of the policies they pursue.

Prospectively, AI might allow autocracies to overcome this disadvantage. The clearest example at present is China (Zeng, 2022), which uses large-scale data collection and AI to support social planning and control (Ding et al., 2020; Pan, 2020)—such as through its Social Credit System (Creemers, 2018; Liang et al., 2018; Sithigh & Siems, 2019). Capitalizing on AI's potential could also help autocracies increase their state capacities through, for example, AI-assisted governance and planning. This in turn could increase the quality of state-provided public services. It also might provide people living in autocracies with greater cultural, economic, or health-related opportunities (Diamandis & Kotler, 2020; Lee & Quifan, 2021).

There are those who might see these benefits as a worthy trade-off with some individual freedoms, leading to strengthened public support for autocracies and state control. Differential opportunities in realizing the potentials of AI might thus reinforce tendencies already evident in countries facing economic, cultural, or security crises (Matovski, 2021). Particularly in times when democracies increasingly find themselves internally challenged with respect to the opportunities with which they provide people, these potentials of AI that asymmetrically favor autocracies represent an obvious challenge to democracies—if realized.

Going further, AI is a technology increasingly discussed in military and security circles (Buchanan & Imbrie, 2022; Goldfarb & Lindsay, 2022). While its normative role and functional potential in these areas are heavily contested, the growing concerns in these circles point to the broad perception that AI could facilitate democracies falling behind autocracies.

Over time, differential trajectories in the development and deployment of AI in democracies and autocracies may emerge. If the assumption holds that autocracies share a greater affinity with AI and can profit more from it than democracies, AI could lead to a power shift between systems and thus weaken democracy.

Artificial Intelligence and Democracy: The Road Ahead

While many AI applications still lie in the future, we already start to see AI's impact on democracy. True, many of AI's future uses and effects remain uncertain. However, it is important that social science engages early on with AI and helps observe, evaluate, and guide its implementation. This includes AI's uses in politics, government, and its regulation and governance. This article provides a conceptual framework that allows the charting of important contact areas between AI and democracy, and connects AI's uses and effects to relevant normative debates in political theory that can provide guidance in the assessment of AI's uses and impacts. Going forward, the broad conceptual contact areas of self-rule, equality, elections, and competition between

systems can serve as topical clusters for future work. Future work will provide a more fine-grained account and advance theories that explain use and effect patterns.

Social scientists need to consider AI in their analysis of features, dangers, and potentials of contemporary democracy. In doing so, they need to reflect the inner workings and domain-specific effects of the underlying technology. At the same time, computer scientists and engineers need to consider the consequences for democracy in AI development and deployment. This means focusing not only on the analysis of the technology itself, but also to consider its embeddedness in economic, political, and social structures that mediate its effects for better or worse. This makes the analysis of AI's impact on democracy an important area of future interdisciplinary work.

The quality of the analysis of AI's effects on democracy depends on specificity regarding the type of AI, how it functions, the conditions for its successful deployment, and the aspect(s) of democracy it touches. Narratives about an unspecified, super-powered AGI and its supposed impact on society may make for stimulating reading but offer little for the analysis of actual effects on society or democracy. In fact, interested parties can use the discussion of AGI and supposed extinction-level event dangers as a smokescreen, distracting public and regulatory attention from more mundane but crucial questions of AI governance, regulation, and the societal distribution of AI-driven gains and risks.

Although AI is often discussed as a danger or threat, it may also provide opportunities to offset some of the contemporary challenges to democracy. Thinking openly about the application of AI in democracy could provide some relief from these challenges. Conscious design choices and transparent audits can help ameliorate dysfunctions and uncover biases.

In general, AI's impact depends on implementation and oversight by the public and regulators. For this, companies, regulators, and society need to be explicit and transparent about what economic, political, or societal goals they want to achieve using AI and how its specific workings can propel or inhibit this pursuit. By nature, this discussion combines normative, mechanistic, and technological arguments and considerations. It is important not to be sidetracked by grandiose, but ultimately imaginary, visions of an AGI, but instead focus on specific instances of narrow AI, their inner workings, uses in specific areas of interest, and effects. This includes the discussion of both potentially positive as well as negative effects.

AI is unlikely to impact many aspects of democracy directly. Nevertheless, public discourse is likely to continue to focus on threats, manipulation, and expected power shifts. This discourse and these expectations have the potential to shape public attitudes toward AI and its impact on democracy strongly, irrespective of their factual basis. Perceived effects can matter more strongly than actual effects. Researchers have a responsibility not to fan the flames of discourse with speculation, but instead remain focused on AI's actual workings and effects.

The broad application of AI in society and politics is only beginning; future developments are, of course, unknown, and so the specific shape AI technology will take, AI's application in society, and its subsequent effects on democracy must be continuously monitored. This article identifies contact areas between AI and democracy that allow for tracking and monitoring this process. More specifically, there are a series of opportunities for future work.

On a very basic level, there is a need for more systematic work on how AI is employed in politics and by governments (Engstrom & Haim, 2023). There are a few early studies—in, for example, the area of predictive policing and the legal system (Brayne, 2021; Ferguson, 2017)—but they are comparatively few and predominantly focused on the United States. There are opportunities for studies of other countries, and for comparative work. Also, much of what we think we know about AI in society is based on journalistic accounts, but while those can be deeply instructive, there is a consistent need for scientific accounts of these developments in different sectors and societies.

There is also a steady stream of studies examining the biases of specific AI systems in particular contexts, such as audits of AI-driven analysis of texts and images and the implementation of AI in societal processes, such as policing, judicial sentencing, credit decisions, or medicine (Bolukbasi et al., 2016; Buolamwini & Gebu, 2018; Caliskan et al., 2017; Mayson, 2019; Mehrabi et al., 2022; S. Mitchell et al., 2021; Obermeyer et al., 2019). However, AI can also be used to study historical biases in society by examining large-scale data sets of historical texts and images to identify the distribution and shifts in the representation of specific societal groups—for example, along gender or racial lines (Jürgens et al., 2022)—or, more generally, identify for further analysis promising cases in which empirical data contradicts model-based expectations (Munk et al., 2022; Rettberg, 2022).

Building on this, systematic work on the regulation of AI in different countries and contexts is also needed (Veale et al., 2023). While these are early days for official regulation of AI, there are already interesting activities—notably, government support for AI-focused research and development. The European Union and the US and Chinese governments have all identified AI as an area of strategic geopolitical and economic competition. Examining competing policy initiatives aimed at driving innovation is a promising early step in examining governments' relationships with AI and related policies and regulations. Earlier comparative studies of the regulation of privacy and platform business models may offer templates for work focused on AI (Bradford, 2023; Cohen, 2019; Farrell & Newman, 2019; Thelen, 2018).

Finally, the development of AI as a scientific and commercial field is a promising subject of research itself. The highly interdisciplinary nature and shifts in the vanguard of AI development between scientific disciplines, academy and business, and geographic areas make this a very interesting case of the contemporary nexus of science, commerce, and society (Lee, 2018; McCorduck, 2004; Metz, 2021). Going

further, AI is increasingly used within the social sciences as a method of discovery. The associated shifts in scientific workflow, questions asked, and approaches to theorizing are a fertile area for future reflection (Grimmer et al., 2021; Mökander & Schroeder, 2022).

There are many promising avenues for future scientific work on the impact of AI on democracy. Here, it is important to combine insights from different fields. Purely technological accounts risk overestimating AI's impact on social systems, given their boundedness and the role of social structures. Accounts coming purely from the social sciences risk misrepresenting the actual workings of existing AI and thereby misattributing its consequences. The conceptual framework presented here offers one approach to combine these interdisciplinary perspectives productively.

The impact of AI on democracy is already progressing. Its systematic, interdisciplinary examination and discussion need to proceed as well.

Acknowledgements

The author thanks Scott Cooper, Valeska Gerstung-Jungherr, Pascal Jürgens, Oliver Posegga, Adrian Rauchfleisch, Ralph Schroeder, Alexander Wuttke, and various anonymous reviewers for their valuable feedback.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research underlying this paper has been generously supported by the *VolkswagenStiftung*.

ORCID iD

Andreas Jungherr  <https://orcid.org/0000-0003-2598-2453>

References

- Acemoglu, D. (2021). Harms of AI. In J. B. Bullock, Y.-C. Chen, J. Himmelreich, V. M. Hudson, A. Korinek, M. M. Young, & B. Zhang (Eds.), *The Oxford handbook of AI governance*. Oxford University Press. <https://doi.org/10.3386/w29247>
- Acemoglu, D., & Johnson, S. (2023). *Power and progress: Our thousand-year struggle over technology and prosperity*. PublicAffairs.
- Acemoglu, D., & Restrepo, P. (2019). Artificial intelligence, automation and work. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The economics of artificial intelligence: An agenda* (pp. 197–236). University of Chicago Press.
- Acemoglu, D., & Restrepo, P. (2022a). Demographics and automation. *The Review of Economic Studies*, 89(1), 1–44. <https://doi.org/10.1093/restud/rdab031>
- Acemoglu, D., & Restrepo, P. (2022b). Tasks, automation, and the rise in U.S. Wage inequality. *Econometrica*, 90(5), 1973–2016. <https://doi.org/10.3982/ECTA19815>

- Achen, C. H., & Bartels, L. M. (2016). *Democracy for realists: Why elections do not produce responsive government*. Princeton University Press.
- Agrawal, A., Gans, J., & Goldfarb, A. (2022). *Prediction machines: The simple economics of artificial intelligence* (Updated and Expanded). Harvard Business Review Press. (Original work published 2018)
- Ahmed, N., Wahed, M., & Thompson, N. C. (2023). The growing influence of industry in AI research. *Science*, 379(6635), 884–886. <https://doi.org/10.1126/science.ade24>
- Asenbaum, H. (2022). Rethinking democratic innovations: A look through the kaleidoscope of democratic theory. *Political Studies Review*, 20(4), 680–690. <https://doi.org/10.1177/14789299211052890>
- Aviram, H., Bragg, A., & Lewis, C. (2017). Felon disenfranchisement. *Annual Review of Law and Social Science*, 13, 295–311. <https://doi.org/10.1146/annurev-lawsocsci-110316-113558>
- Bai, H., Voelkel, J. G., Eichstaedt, J. C., & Willer, R. (2023). Artificial intelligence can persuade humans on political issues. *OSF Preprints*. <https://doi.org/10.31219/osf.io/stakv>
- Bakhtin, A., Brown, N., Dinan, E., Farina, G., Flaherty, C., Fried, D., Goff, A., Gray, J., Hu, H., Jacob, A. P., Komeili, M., Konath, K., Kwon, M., Lerer, A., Lewis, M., Miller, A. H., Mitts, S., Renduchintala, A., & Zijlstra, M., Meta Fundamental AI Research Diplomacy Team. (2022). Human-level play in the game of Diplomacy by combining language models with strategic reasoning. *Science*, 378(6624), 1067–1074. <https://doi.org/10.1126/science.ade9097>
- Bennett, W. L., & Livingston, S. (Eds.). (2021). *The disinformation age: Politics, technology, and disruptive communication in the United States*. Cambridge University Press. <https://doi.org/10.1017/9781108914628>
- Bessen, J. (2022). *New goliaths: How corporations use software to dominate industries, kill innovation, and undermine regulation*. Yale University Press.
- Bimber, B. (2003). *Information and American democracy: Technology in the evolution of political power*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511615573>
- Bisbee, J., Clinton, J., Dorff, C., Kenkel, B., & Larson, J. (2023). Artificially precise extremism: How internet-trained LLMs exaggerate our differences. *Socarxiv*. <https://doi.org/10.31235/osf.io/5ecfa>
- Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., & Kalai, A. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In D. D. Lee, U. von Luxburg, R. Garnett, M. Sugiyama, & I. Guyon (Eds.), *NIPS'16: Proceedings of the 30th international conference on neural information processing systems* (pp. 4356–4364). Curran Associates Inc. <https://doi.org/10.5555/3157382.3157584>
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Bradford, A. (2023). *Digital empires: The global battle to regulate technology*. Oxford University Press.
- Brayne, S. (2021). *Predict and surveil: Data, discretion, and the future of policing*. Oxford University Press. <https://doi.org/10.1093/oso/9780190684099.001.0001>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., & Amodei, D. (2020). Language models are few-shot learners. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), *Advances in neural information processing systems* (Vol. 33, pp. 1877–1901). Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf>
- Brynjolfsson, E., Jin, W., & Wang, X. (2023). Information technology, firm size, and industrial concentration. *NBER Working Paper*, 31065. <https://doi.org/10.3386/w31065>
- Brynjolfsson, E., & McAfee, A. (2016). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W.W. Norton.
- Buchanan, B., & Imbrie, A. (2022). *The new fire: War, peace, and democracy in the age of AI*. MIT Press.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In S. A. Friedler & C. Wilson (Eds.), *Proceedings of the 1st conference on fairness, accountability and transparency* (Vol. 81, pp. 77–91). Proceedings of Machine Learning Research (PMLR). <https://proceedings.mlr.press/v81/buolamwini18a.html>
- Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183–186. <https://doi.org/10.1126/science.aal4230>
- Castells, M. (2013). *Communication power* (2nd ed.). Oxford University Press. (Original work published 2009)
- Cho, W. K. T., & Cain, B. E. (2020). Human-centered redistricting automation in the age of AI. *Science*, 369(6508), 1179–1181. <https://doi.org/10.1126/science.abd1879>
- Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data*, 5(2), 153–163. <https://doi.org/10.1089/big.2016.0047>
- Chow, S., Liver, S., & Nelson, A. (2018). Streamlining bioactive molecular discovery through integration and automation. *Nature Reviews Chemistry*, 2, 174–183. <https://doi.org/10.1038/s41570-018-0025-7>
- Christian, B. (2020). *The alignment problem: Machine learning and human values*. W.W. Norton.
- Cohen, J. E. (2019). *Between truth and power: The legal constructions of informational capitalism*. Oxford University Press. <https://doi.org/10.1093/oso/9780190246693.001.0001>
- Converse, P. (1964). The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and discontent* (pp. 206–261). Free Press.
- Creemers, R. (2018). *China's social credit system: An evolving practice of control*. Social Science Research Network (SSRN). <https://doi.org/10.2139/ssrn.3175792>
- Criddle, C., & Murphy, H. (2023). OpenAI chief says new rules are needed to guard against AI risks. *Financial Times*. <https://www.ft.com/content/aa3598f7-1470-45e4-a296-bd26953c176f>
- Dahl, R. A. (1998). *On democracy*. Yale University Press.
- Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press.
- Diamandis, P. H., & Kotler, S. (2020). *The future is faster than you think: How converging technologies are transforming business, industries, and our lives*. Simon & Schuster.
- Ding, J., Chun, A., Liu, Y.-L., Han, E., Lewis, D., Gal, D., & Creemers, R. (2020). *The AI powered state: China's approach to public sector innovation*. Nesta. <https://apo.org.au/sites/default/files/resource-files/2020-05/apo-nid305076.pdf>
- Douek, E. (2021). Governing online speech: From “posts-as-trumps” to proportionality and probability. *Columbia Law Review*, 121(3), 759–834. <https://doi.org/10.2139/ssrn.3679607>

- Eisenstein, E. L. (1979). *The printing press as an agent of change*. Cambridge University Press. <https://doi.org/10.1017/CBO9781107049963>
- Engstrom, D. F., & Haim, A. (2023). Regulating government AI and the challenge of sociotechnical design. *Annual Review of Law and Social Science*, 19, 091626. <https://doi.org/10.1146/annurev-lawsocsci-120522-091626>
- Espinoza, J., & Johnston, I. (2023). EU lawmakers agree tough measures over use of AI. *Financial Times*. <https://www.ft.com/content/da597e19-4d63-4d4d-b7d1-c3405302a1e3>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Farrell, H., & Newman, A. L. (2019). *Of privacy and power: The transatlantic struggle over freedom and security*. Princeton University Press.
- Ferguson, A. G. (2017). *The rise of big data policing: Surveillance, race, and the future of law enforcement*. New York University Press.
- Filgueiras, F. (2022). The politics of AI: Democracy and authoritarianism in developing countries. *Journal of Information Technology & Politics*, 19(4), 449–464. <https://doi.org/10.1080/19331681.2021.2016543>
- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80(1), 298–320. <https://doi.org/10.1093/poq/nfw006>
- Frey, C. B. (2019). *Technology trap: Capital, labor, and power in the age of automation*. Princeton University Press.
- Gallego, A., & Kurer, T. (2022). Automation, digitalization, and artificial intelligence in the workplace: Implications for political behavior. *Annual Review of Political Science*, 25, 463–484. <https://doi.org/10.1146/annurev-polisci-051120-104535>
- Gellers, J. C., & Gunkel, D. J. (2023). Artificial intelligence and international human rights law: Implications for humans and technology in the 21st century and beyond. In A. Zwitter & O. J. Gstrein (Eds.), *Handbook on the politics and governance of big data and artificial intelligence* (pp. 430–455). Edward Elgar Publishing. <https://doi.org/10.4337/9781800887374.00027>
- Goldfarb, A., & Lindsay, J. R. (2022). Prediction and judgment: Why artificial intelligence increases the importance of humans in war. *International Security*, 46(3), 7–50. https://doi.org/10.1162/isec_a_00425
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Grimmer, J., Roberts, M. E., & Stewart, B. M. (2021). Machine learning for social science: An agnostic approach. *Annual Review of Political Science*, 24, 395–419. <https://doi.org/10.1146/annurev-polisci-053119-015921>
- Guttman, A. (2007). Democracy. In R. E. Goodin, P. Pettit, & T. Pogge (Eds.), *A companion to contemporary political philosophy* (2nd ed., pp. 521–532). Blackwell. <https://doi.org/10.1002/9781405177245.ch25>
- Hanson, R. (2016). *The age of Em: Work, love, and life when robots rule the earth*. Oxford University Press. <https://doi.org/10.1093/oso/9780198754626.001.0001>
- Healy, K. (2017). Fuck nuance. *Sociological Theory*, 35(2), 118–127. <https://doi.org/10.1177/0735275117709046>
- Hersh, E. D. (2015). *Hacking the electorate: How campaigns perceive voters*. Cambridge University Press. <https://doi.org/10.1017/CBO9781316212783>
- Horton, J. J. (2023). *Large language models as simulated economic agents: What can we learn from Homo Silicus?* [NBER Working Papers]. https://www.nber.org/system/files/working_papers/w31122/w31122.pdf
- Issenberg, S. (2012). *The victory lab: The secret science of winning campaigns*. Crown Publishing Group.
- Jungherr, A., Rivero, G., & Gayo-Avello, D. (2020). *Retooling politics: How digital media are shaping democracy*. Cambridge University Press. <https://doi.org/10.1017/9781108297820>
- Jungherr, A., & Schroeder, R. (2022). *Digital transformations of the public arena*. Cambridge University Press. <https://doi.org/10.1017/9781009064484>
- Jungherr, A., & Schroeder, R. (2023). Artificial intelligence and the public arena. *Communication Theory*, 33(3), qtad006. <https://doi.org/10.1093/ct/qtad006>
- Jungherr, A., Schroeder, R., & Stier, S. (2019). Digital media and the surge of political outsiders: Explaining the success of political challengers in the United States, Germany, and China. *Social Media + Society*, 5(3), 1–12. <https://doi.org/10.1177/2056305119875439>
- Jürgens, P., Meltzer, C. E., & Scharkow, M. (2022). Age and gender representation on German TV: A longitudinal computational analysis. *Computational Communication Science*, 4(1), 5. <https://doi.org/10.5117/CCR2022.1.005.JURG>
- Kaufmann, T. (2019). *Die Mitte der Reformation: Eine Studie zu Buchdruck und Publizistik im deutschen Sprachgebiet, zu ihren Akteuren und deren Strategien, Inszenierungs- und Ausdrucksformen* [The middle of the Reformation: A study on book printing and journalism in the German-speaking area, on their actors and their strategies, staging and forms of expression]. Mohr Siebeck. <https://doi.org/10.1628/978-3-16-156606-6>
- Kaye, D. (2018). *Report on artificial intelligence technologies and implications for freedom of expression and the information environment*. United Nations Human Rights Office of the High Commissioner. <https://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/ReportGA73.aspx>
- Keane, J. (2013). *Democracy and media decadence*. Cambridge University Press. <https://doi.org/10.1017/CBO9781107300767>
- Kim, J., & Lee, B. (2023). AI-augmented surveys: Leveraging large language models for opinion prediction in nationally representative surveys. *Arxiv*. <https://arxiv.org/abs/2305.09620>
- King, R. D., Rowland, J., Oliver, S. G., Young, M., Aubrey, W., Byrne, E., Liakata, M., Markham, M., Pir, P., Soldatova, L. N., Sparkes, A., Whelan, K. E., Whelan, K. E., & Clare, A. (2009). The automation of science. *Science*, 324(5923), 88–89. <https://doi.org/10.1126/science.1165620>
- Kitchens, B., Johnson, S. L., & Gray, P. (2020). Understanding echo chambers and filter bubbles: The impact of social media on diversification and partisan shifts in news consumption. *MIS Quarterly*, 44(4), 1619–1649. <https://doi.org/10.25300/MISQ/2020/16371>
- Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 1–12. <https://doi.org/10.1177/2053951714528481>
- Krebs, S., McCain, R. M., & Brundage, M. (2022). All the news that's fit to fabricate: AI-generated text as a tool of media misinformation. *Journal of Experimental Political Science*, 9(1), 104–117. <https://doi.org/10.1017/XPS.2020.37>
- Kreiss, D. (2016). *Prototype politics: Technology-intensive campaigning and the data of democracy*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199350247.001.0001>

- Kuklinski, J. H., & Quirk, P. J. (2000). Reconsidering the rational public: Cognition, heuristics, and mass opinion. In A. Lupia, M. D. McCubbins, & S. L. Popkin (Eds.), *Elements of reason: Cognition, choice, and the bounds of rationality* (pp. 153–182). Cambridge University Press.
- Kuran, T. (1995). *Private truths, public lies: The social consequences of preference falsification*. Harvard University Press.
- Landemore, H. (2012). *Democratic reason: Politics, collective intelligence, and the rule of the many*. Princeton University Press.
- Landemore, H. (2022). *Can AI bring deliberative democracy to the masses?* [Working Paper]. <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>
- Landemore, H., & Elster, J. (Eds.). (2012). *Collective wisdom: Principles and mechanisms*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511846427>
- Larson, E. J. (2021). *The myth of artificial intelligence: Why computers can't think the way we do*. The Belknap Press of Harvard University.
- Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: Traps in big data analysis. *Science*, 343(6176), 1203–1205. <https://doi.org/10.1126/science.1248506>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. <https://doi.org/10.1038/nature14539>
- Lee, K.-F. (2018). *AI superpowers: China, Silicon Valley, and the new world order*. Houghton Mifflin Harcourt.
- Lee, K.-F., & Quifan, C. (2021). *AI 2041: Ten visions for our future*. Currency.
- Liang, F., Das, V., Kostyuk, N., & Hussain, M. M. (2018). Constructing a data-driven society: China's social credit system as a state surveillance infrastructure. *Policy & Internet*, 10(4), 415–453. <https://doi.org/10.1002/poi3.183>
- Lindblom, C. E. (1965). *Intelligence of democracy: Decision making through mutual adjustment*. Free Press.
- Lindblom, C. E. (2001). *The market system: What it is, how it works, and what to make of it*. Yale University Press.
- Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139032490>
- Lupia, A., & McCubbins, M. D. (1998). *The democratic dilemma: Can citizens learn what they need to know?* Cambridge University Press.
- Matovski, A. (2021). *Popular dictatorships: Crises, mass opinion, and the rise of electoral authoritarianism*. Cambridge University Press. <https://doi.org/10.1017/9781009047500>
- Mayson, S. G. (2019). Bias in, bias out. *The Yale Law Journal*, 128(8), 2218–2300.
- McCorduck, P. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. A K Peters.
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2022). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35. <https://doi.org/10.1145/3457607>
- Metz, C. (2021). *Genius makers: The mavericks who brought AI to Google, Facebook, and the world*. Dutton.
- Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Farrar, Straus and Giroux.
- Mitchell, S., Potash, E., Barocas, S., D'Amour, A., & Lum, K. (2021). Algorithmic fairness: Choices, assumptions, and definitions. *Annual Review of Statistics and Its Application*, 8, 141–163. <https://doi.org/10.1146/annurev-statistics-042720-125902>
- Mökander, J., & Schroeder, R. (2022). AI and social theory. *AI & Society*, 37, 1337–1351. <https://doi.org/10.1007/s00146-021-01222-z>
- Müller, J.-W. (2021). *Democracy rules*. Allen Lane.
- Munk, A. K., Olesen, A. G., & Jacomy, M. (2022). The thick machine: Anthropological AI between explanation and explanation. *Big Data & Society*, 9(1), 1–14. <https://doi.org/10.1177/20539517211069891>
- Narayanan, A. (2023). *Understanding social media recommendation algorithms*. Knight First Amendment Institute at Columbia University. <https://knightcolumbia.org/content/understanding-social-media-recommendation-algorithms>
- Neuman, W. R. (1991). *The future of the mass audience*. Cambridge University Press.
- Nickerson, D. W., & Rogers, T. (2014). Political campaigns and big data. *The Journal of Economic Perspectives*, 28(2), 51–74. <https://doi.org/10.1257/jep.28.2.51>
- Nielsen, R. K. (2020). Economic contexts of journalism. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The handbook of journalism studies* (2nd ed., pp. 324–340). Routledge. <https://doi.org/10.4324/9781315167497-21>
- Nilsson, N. J. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511819346>
- Ober, J. (2008). *Democracy and knowledge: Innovation and learning in classical Athens*. Princeton University Press.
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>
- Pan, J. (2020). *Welfare for autocrats: How social assistance in China cares for its rulers*. Oxford University Press. <https://doi.org/10.1093/oso/9780190087425.001.0001>
- Papacharissi, Z. (2021). *After democracy: Imagining our political future*. Yale University Press.
- Parmar, N. J., Vaswani, A., Uszkoreit, J., Kaiser, L., Shazeer, N., Ku, A., & Tran, D. (2018). Image transformer. In J. Dy & A. Krause (Eds.), *Proceedings of the 35th international conference on machine learning* (pp. 4055–4064). PMLR. <http://proceedings.mlr.press/v80/parmar18a/parmar18a.pdf>
- Pearl, J. (2019). The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, 62(3), 54–60. <https://doi.org/10.1145/3241036>
- Phillips, A. (2021). *Unconditional equals*. Princeton University Press.
- Popkin, S. L. (1991). *The reasoning voter: Communication and persuasion in presidential campaigns*. The University of Chicago Press.
- Prior, M. (2007). *Post-broadcast democracy: How media choice increases inequality in political involvement and polarizes elections*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139878425>
- Prior, M. (2017). Conditions for political accountability in a high-choice media environment. In K. Kenski & K. H. Jamieson (Eds.), *The Oxford handbook of political communication* (pp. 897–912). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199793471.013.63>
- Prior, M. (2018). *Hooked: How politics captures people's interest*. Cambridge University Press.

- Przeworski, A. (1991). *Democracy and the market: Political and economic reforms in Eastern Europe and Latin America*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139172493>
- Przeworski, A. (2018). *Why bother with elections?* Polity Press.
- Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical text-conditional image generation with CLIP latents. *Arxiv*. <https://www.crcv.ucf.edu/wp-content/uploads/2018/11/Group-6-Paper-2-Dalle-2.pdf>
- Rettberg, J. W. (2022). Algorithmic failure as a humanities methodology: Machine learning's mispredictions identify rich cases for qualitative analysis. *Big Data & Society*, 9(2), 1–6. <https://doi.org/10.1177/20539517221131290>
- Risse, M. (2023). *Political theory of the digital age: Where artificial intelligence might take us*. Cambridge University Press. <https://doi.org/10.1017/9781009255189>
- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson Education. (Original work published 1995).
- Sithigh, D. M., & Siems, M. (2019). The Chinese Social Credit System: A model for other countries? *Morden Law Review*, 82(6), 1034–1071. <https://doi.org/10.1111/1468-2230.12462>
- Sanders, N. E., & Schneier, B. (2021). Machine learning featurizations for AI hacking of political systems. *Arxiv*. <https://arxiv.org/pdf/2110.09231.pdf>
- Scharkow, M., Mangold, F., Stier, S., & Breuer, J. (2020). How social network sites and other online intermediaries increase exposure to news. *Proceedings of the National Academy of Sciences of the United States of America*, 117(6), 2761–2763. <https://doi.org/10.1073/pnas.1918279117>
- Schlozman, K. L., Brady, H. E., & Verba, S. (2018). *Unequal and unrepresented: Political inequality and the people's voice in the new gilded age*. Princeton University Press.
- Schneider, G. (2018). Automating drug discovery. *Nature Reviews Drug Discovery*, 17, 97–113. <https://doi.org/10.1038/nrd.2017.232>
- Schölkopf, B., Locatello, F., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A., & Bengio, Y. (2021). Toward causal representation learning. *Proceedings of IEEE*, 109(5), 612–634. <https://doi.org/10.1109/JPROC.2021.3058954>
- Schudson, M. (1978). *Discovering the news: A social history of American newspapers*. Basic Books.
- Schwartzberg, M. (2015). Epistemic democracy and its challenges. *Annual Review of Political Science*, 18, 187–203. <https://doi.org/10.1146/annurev-polisci-110113-121908>
- Settle, J. E. (2018). *Frenemies: How social media polarizes America*. Cambridge University Press. <https://doi.org/10.1017/9781108560573>
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529, 484–489. <https://doi.org/10.1038/nature16961>
- Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, 362(6419), 1140–1144. <https://doi.org/10.1126/science.aar6404>
- Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., & Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, 550, 354–359. <https://doi.org/10.1038/nature24270>
- Smith, B. C. (2019). *The promise of artificial intelligence: Reckoning and judgment*. MIT Press.
- Stasavage, D. (2020). *The decline and rise of democracy: A global history from antiquity to today*. Princeton University Press.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.
- Thelen, K. (2018). Regulating uber: The politics of the platform economy in Europe and the United States. *Perspectives on Politics*, 16(4), 938–953. <https://doi.org/10.1017/S1537592718001081>
- Tilly, C. (2007). *Democracy*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511804922>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. V. N. Vishwanathan, & R. Garnett (Eds.), *NIPS 2017: 31st conference on neural information processing systems* (Vol.30, pp. 1–11). Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547de91fbd053c1c4a845aa-Paper.pdf>
- Veale, M., Matus, K., & Gorwa, R. (2023). AI and global governance: Modalities, rationales, tensions. *Annual Review of Law and Social Science*, 19, 040749.
- Wilkinson, A. (2023). The looming threat of AI to Hollywood, and why it should matter to you. *Vox*. <https://www.vox.com/culture/23700519/writers-strike-ai-2023-wga>
- Williams, B. A., & Carpini, M. X. D. (2011). *After broadcast news: Media regimes, democracy, and the new information environment*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511846366>
- Winner, L. (1980). Do artifacts have politics? *Daedalus*, 109(1), 121–136.
- Wintrobe, R. (1998). *The political economy of dictatorship*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139174916>
- Young, I. M. (2002). *Inclusion and democracy*. Oxford University Press. <https://doi.org/10.1093/0198297556.001.0001>
- Zaller, J. R. (1992). *The nature and origins of mass opinion*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511818691>
- Zeng, J. (2022). *Artificial intelligence with Chinese characteristics: National strategy, security and authoritarian governance*. Palgrave Macmillan. <https://doi.org/10.1007/978-981-19-0722-7>

Author Biography

Andreas Jungherr (Dr. rer. pol. University of Bamberg) holds the Chair for Political Science, especially Digital Transformation at the University of Bamberg. His research interests include digital media in politics and society, political communication, and computational social science.