Introduction

Artificial Intelligence, Politics, and Political Science

Nathaniel Persily & Joshua A. Tucker

Introduction

Artificial Intelligence is reshaping politics and political science, just as it is transforming other social phenomena and their associated academic fields. As with the larger policy debates over artificial intelligence and its social impact, attitudes toward the newest version of this technology range from utopian to dystopian, with many also alleging the technology is overhyped, at least in the short term. With the pace of technological and political change nearly outpacing the capacity of (human) academics to analyze these trends, any endeavor to take stock of where things stand for AI and politics in the summer of 2026 is necessarily fraught. Although the trajectory remains uncertain, a volume like the one that follows provides a critical snapshot of the state of the field as it begins to grapple with the multifaceted questions of AI's relationship to politics and research.

It was against this background, in the fall of 2025, that Taeku Lee, then President of the American Political Science Association, commissioned a Presidential Task Force on AI, Politics, and Political Science, naming us as co-chairs. We then identified the topics that would comprise the report and identified potential committee chairs for each topic. Upon agreeing to serve on the task force, each committee chair then assembled committees to write the chapters in this volume, which collectively represent the "Report" of the Task Force. Admittedly, as an edited volume in which different authors express disagreements and different perspectives on the themes related to AI and politics, this Report differs from other task force reports. However, with (humble) recognition of the variety of views on the relevant questions at this early stage of technological development, we thought that surfacing these disagreements was itself a form of service to the discipline. Moreover, the topic is so vast, multifaceted, and quickly changing that crafting a volume touching on many, if admittedly not all, the subfields of political science in the space of six months required that we grant considerable autonomy to each committee. In that spirit, members of the Task Force speak only for themselves in the chapters that they have authored, and the volume represents a collection of views rather than a definitive assessment of the appropriate place of AI in the study and practice of politics.

Our goal with this volume, therefore, is not to offer predictions about the future of AI but, rather, to map the state of existing research, identify the most pressing questions, and establish a

foundation for the rigorous empirical study of AI's effects on politics and governance. None of this is to suggest that we must resign ourselves to mere speculation in assessing the impact of artificial intelligence: AI is already affecting the practice and study of politics in identifiable ways. Regulatory bodies throughout the United States and around the world have already begun to grapple with the use of generative AI in political campaigning, for instance. Administrative agencies are rolling out AI tools to deal with case backlogs and to improve public services. The issue of AI in warfare is becoming increasingly relevant in armed conflicts in Ukraine, the Middle East, and around the world. And political scientists, like other scientists and social scientists, are employing new AI tools to study a range of political phenomena – so much so, in fact, that we are already seeing glimpses of a transformed profession in which researcher productivity outstrips the capacity of traditional institutions (such as peer-reviewed journals) to curate and to present political science to a larger audience.[1]

The Scope of This Volume

One of the challenges with a volume like this is that the topic is potentially so capacious that disagreements inevitably emerge as to what phenomena are in and out of scope. Different people mean different things when they reference AI, and the range of topics within political science that might be relevant to AI arguably covers the entire discipline. We hope that this volume begins a conversation about AI and the discipline of political science and that omissions are filled in by later work.

We begin with the question as to what we mean by artificial intelligence. At its most general and abstract, the concept could include all forms of machine learning and digital algorithms, but surely that definition is both overbroad and underinclusive. In current popular imagination, largely shaped by consumer experiences with ChatGPT, generative AI (in the form of chatbots and image generators) has been the primary focus of any discussion concerning AI and politics, especially as it pertains to issues related to deep fakes and the generation of AI content on social media. However, generative AI is just one incarnation of AI (let alone politically relevant AI), the definition of which must also include the brave new world of agentic systems and "embodied AI" in the form of robotics, autonomous vehicles, and drones. In the national security context, for example, these latter types of AI may be most significant.

The definition of AI eventually adopted by the European Union in its AI Act is a typical one: "a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or

[1] See Munger, K., Bakker, B. N., Berinsky, A. J., Just, N., Guess, A. M., Giger, N., Tenenboim-Weinblatt, K., Lawrence, R., & van de Rijt, A. (2026). *Peer review 2027: Scenarios for academic publishing in the age of AI* [Preprint]. SocArXiv. https://osf.io/preprints/socarxiv/594zj_v1

decisions that can influence physical or virtual environments."[2] However, the EU arrived at that definition only after several years of debate and after having gone back to the drawing board following the explosion of generative systems following the launch of ChatGPT. This definition is as good as any, so long as we keep in mind that the world of politically relevant AI includes all of the related systems (robotics etc.) described above. For purposes of this volume, we do not feel the need to settle any debate over the proper definition of AI, but, rather, we want to emphasize that it includes a family of technologies that are ever-evolving and that focusing only on the most popular AI products today (such as chatbots) ignores a diverse array of tools with political relevance.

As difficult as it may be to define AI, it is equally challenging to fence off the terrain of politics and political science that might be relevant to, or affected by, AI. In a sense, any of AI's large-scale social benefits or harms could properly be within the scope of "politics." To the extent AI poses existential risks – such as through increased probabilities of the development of rogue chemical, biological, radiological, or nuclear (so-called CBRN risk) weapons or through acceleration of climate change, massive labor displacement, or a Skynet-Terminator scenario of robot domination of humans – all those developments are properly described as "political." Similarly, if AI helps cure diseases like cancer, dramatically increases life expectancy and the quality of life, or leads to massive economic growth enabling millions to lead a life of leisure, politics will be affected.

These potential impacts go far beyond the heuristics that many bring to thinking about AI and politics. The mental model many analysts bring to these questions was developed in the social media age, when technology-related concerns about misinformation, polarization, privacy violations, and political manipulation were top of mind. This volume deals with those issues, but, as argued later in this introduction, we worry that an excessive focus on AI's relationship to social media understates the relevant impact of AI on politics. In contrast, we have attempted to include in this volume representatives from as many different subfields of political science as is reasonably possible. To be sure, not every subfield of political science is represented, but we hope that this book provides a springboard for similar volumes, perhaps each focusing on a particular subdiscipline of political science.

Finally, this volume has the dual ambition of not merely describing the politics of AI but also reflecting on what AI means for the practice of political science. Here, too, we do not cover all possible AI-related impacts on the job of a political scientist. Indeed, depending on how capable AI systems become, we can only begin to speculate what these technological developments might mean for the future of universities or the role of humans as researchers and teachers (let alone the expectations of students as to what is worth learning in the AI age). The chapters that focus on the profession consider the impact of AI on traditional forms of political science research, writing, publishing, and teaching: the methods and tools we use to perform that job as it has evolved over

[2] See the EU AI Act, Article 3 (1). https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng?utm_source=chatgpt.com

the last century or so. We recognize, however, that the tectonic plates underlying university research and teaching are shifting, and that in a few decades, we might look back on this volume with a nostalgia for the traditional job of a political scientist, which AI changed into a fundamentally different occupation.

But the stakes here are not just professional. To the extent that political science provides a framework for societies to address distributions of power with shared evidence and methods, it is incumbent upon political science as a discipline to wrestle with the way AI will change our *standards* of methods and evidence. If AI fundamentally changes the way we produce knowledge as a field – which is still a big if – then the quality of the arguments we can present to the public may change as well. Understanding the nature of those changes is another goal of this volume.

Temporal, Geographic, and Data Challenges in AI Research

We would be remiss not to acknowledge the significant challenges that attend any effort to assess the influence of a fast-developing technology with global implications on a broad array of political phenomena and the academic study of them. Even as scholars who frequently write about technology and politics, we find it particularly challenging to keep up with the latest AI developments, let alone to publish analyses of their political implications in a timely manner.

Between the time our contributors drafted their chapters and the time this volume reaches readers, new models have been released, new capabilities have emerged, and some of the specific examples cited here may already feel dated. The lag between technological development and scientific analysis of its impact is an unavoidable feature of the landscape. For example, after the authors submitted their chapters, we have witnessed dramatic developments in the deployment of agentic AI systems (such as Claude Code and Codex) with implications for political science research, let alone the larger economy. Over the past year, we also have seen new controversies, such as the one involving Anthropic and the U.S. Department of War, that seem to present unique and fundamental questions related to the "politics of AI," let alone the spectre presented by Anthropic's Mythos model of a product that might allow a company to compromise critical systems for entire nation states. Several of our contributors note the difficulty of writing about a technology with capabilities that shift on the order of weeks and months rather than years. We have, therefore, tried to focus the volume on analytical frameworks, conceptual distinctions, and empirical findings that will retain their value even as the specific technologies evolve.

A second challenge is the geography of AI development and concomitant research on AI. The production of frontier AI systems is overwhelmingly concentrated in the United States and China, with significant but secondary contributions from a handful of other countries. This concentration shapes the research in predictable ways. Much of the empirical literature on AI and politics is based on the American context, and the dominant AI models are trained primarily on English-language and Chinese-language data. The implications of this linguistic and geographic

asymmetry are a recurring theme in the chapters that follow. AI models may perform differently in low-resource languages. For example, they may hallucinate more frequently, fail to capture cultural nuances, or answer politically sensitive questions differently. The risk that the information environments of the Global South will be polluted with lower-quality AI-generated content is real and underappreciated.

Similarly, to the extent that this volume focuses on the existing research on AI and politics, which has a pronounced Western bias, it necessarily leaves out much of the critical work done on these topics in the rest of the world. This volume, incomplete though it may be, should also be read as a clarion call for a geographic broadening of the relevant research agenda. To the extent that AI "lowers the cost" of politics – i.e., by reducing the costs of carrying out fundamental political activities, from running political campaigns to surveilling political opponents – we might expect its impact to be felt even more profoundly in developing countries. The field of political science needs to account for the global impact of this technology and dedicate scholarly resources to ensure an inclusive assessment of the far-reaching political impacts of AI.

A third challenge, and one that will be familiar to readers of our earlier work on social media (Persily and Tucker, 2020), concerns the problem of data access. The training data, model architectures, and user interaction logs that would be most valuable for studying AI's political effects are almost entirely locked behind corporate firewalls. We are, once again, in a position where a small number of extraordinarily powerful private firms control access to the data necessary to assess the impact of a new technology on society. This is not to say we do not have tools available to study AI; as we explain in more detail below, we may actually be in a better place, due to the business models of AI firms, to observe the outputs of these models than we were for, say, social media algorithms. But we remain hamstrung in our ability to observe how human beings interact with these AI systems – to say nothing of the inner workings of the systems themselves.

Lessons from the Study of Social Media

When analyzing the likely effect of AI on politics, commentators naturally gravitate to the most recent example where technological innovation has been blamed for political change: social media. The reliance on this paradigm is unsurprising, as some of the most ardent criticisms of AI focus on its nexus to social media, through generation and propagation of deep fakes or hate speech, or other familiar dangers related to privacy, children, addiction, or emotional well-being. However, AI is a general purpose technology with implications for the economy and society that far exceed even the direst warnings (or highest praise) about social media. For all of the heated criticisms, no one ever feared that social media would lead to mass unemployment or human extinction. Although we agree that the rise of social media can be a useful heuristic for thinking about AI, it is important that we do not simply default to lessons learned about social media and assume things will be the same with AI. Wholesale transference of the political anxieties developed

with respect to social media misrepresents the potential, for both good and ill, of this newest family of technologies to have a substantial impact on politics and governance.

The levels of skepticism and concern Americans express relating to AI are, no doubt, a product of the unrelenting criticism (both warranted and unwarranted) of social media over the past decade. It should be noted that in the early days of social media, the technology was hailed as a "liberation technology" and lauded for its likely democratizing impacts (Diamond, 2010). However, over the past decade, Americans have come to distrust both the government and Big Tech. Put those two together and it is no surprise that large majorities express concern about AI and feel that it will do more harm than good.[3] Those anxieties have now led to explicit acts of political violence, as anti-AI activists threw a molotov cocktail at the house of OpenAI CEO Sam Altman and fired gun shots at the home of an Indianapolis councilman with a note placed on his doorstep saying "no data centers." If the frame that people bring to these questions is one forged from worries, for example, that social media caused untold political harm around the globe, we should not be surprised by the level of political anxiety accompanying the rollout of AI. Nor should we be surprised that most other countries, with populations expressing more favorable opinions of social media, also have more favorable views about AI.[4]

In this section, we consider four characteristics of social media research in political science to briefly assess the extent to which we think these prior experiences with social media are or are not likely to inform our understanding of AI and its relationship to the study of politics. These are the ways in which the new technology (formerly social media, now AI) (1) plugs into existing research topics in political science; (2) creates new methodological challenges; (3) drives methodological innovation; and (4) creates new challenges related to data access.

Plugging into existing political science research topics

In addition to functioning as a new source of data, social media also increasingly attracted attention from political scientists because of its role as a potentially important independent variable affecting important outcomes of interest in political science. Social media originally burst into the collective consciousness of scholars studying protest, first in authoritarian regimes (e.g., Arab Spring) and then later in democracies (e.g., Occupy Wall Street). However, it was not long before the number of political topics where the study of social media seemed relevant grew substantially, coming

[3] Quinnipiac University. (2026, March 30). *The age of artificial intelligence: Americans' AI use increases while views on it sour* [Poll]. https://poll.qu.edu/images/polling/us/us03302026_uaio53.pdf

[4] See Wike, R., Silver, L., Fetterolf, J., Huang, C., Austin, S., Clancy, L., & Gubbala, S. (2022, December 6). *Views of social media and its impacts on society in advanced economies*. Pew Research Center. https://www.pewresearch.org/global/2022/12/06/views-of-social-media-and-its-impacts-on-society-in-advanced-economies-2022/; Capstick, E. (2025). Public opinion. In *AI index report 2025* (Chap. 8). Stanford Institute for Human-Centered Artificial Intelligence. https://hai.stanford.edu/assets/files/hai_ai-index-report-2025_chapter8_final.pdf

over time to include voting behavior, public opinion formation, elite communication and campaigning, foreign influence campaigns and, perhaps most significantly, political polarization and the political information environment. As this volume demonstrates, AI seems poised to follow a similar trajectory to social media in this regard, as we suspect political scientists will increasingly find themselves wrestling with substantive questions related to AI. Although the direct effect of artificial intelligence on the political topics of interest may not be quite as clear as was the case for social media, the wide range of topics covered in this volume illustrates that AI may prove in the long term to be even more consequential in terms of the sheer number of political arenas (e.g., defense, courts, and procurement in addition to all of the more obvious information environment implications) in which its impact may be felt.

Methodological challenges

Social media data did not look like traditional political science data in so far as they did not fit neatly into a spreadsheet where all of the cells were filled with numbers. Instead, social media contained text (and eventually images and video) in a networked format, supplemented by multiple levels of metadata about posts, users, content, devices, etc. In many ways, the rise of AI and, in particular, generative AI, simply continues these trends, with text, images, and videos increasingly serving as the units of analysis. Another methodological challenge to studying social media was that the subject of study itself was changing more rapidly than other political subjects of study. Kevin Munger coined the term "temporal validity" to address the question of whether the conclusions from a study of, for example, Facebook in 2018 would still be considered valid for understanding Facebook's impact in 2020.[5] The study of AI faces a similar, and potentially even more extreme, challenge in this regard, as the underlying models that are driving the outputs from AI systems seem to be changing at an even faster rate than social media platforms and algorithms. Moreover, the differences in technological capabilities from one frontier model to the next feel much more significant, especially recently, than the changes from one style of algorithmic feed to the next on a social media platform.

Methodological Innovation

The advent of social media went hand in hand with the text-as-data revolution in the field, as high performance computing ushered in an era when rigorous quantitative analysis was now possible using text in addition to numbers. Although not limited to social media data alone, the challenges of working with these data led to political scientists developing skills in machine learning and network analysis. Here perhaps, the difference between social media and AI is most stark. Whereas social media provided new forms of data that spurred methodological innovation, AI itself is a

[5] Munger, K. (2023). Temporal validity as meta-science. *Research & Politics, 10*(3). https://journals.sagepub.com/doi/10.1177/20531680231187271

form of methodological innovation. While drawing the line between machine learning and what is now generically called AI is an imprecise exercise – and, as noted above, one in which we have deliberately chosen not to engage in this volume – it is clear that AI, and especially the more recently emerging agentic AI tools, such as Claude Code or Open AI's Codex, have the potential to radically change many aspects of the research process in which political scientists engage. These changes may be orders of magnitude more dramatic – and may happen much more rapidly – than in the past.[6]

The Need for Data Access

Finally, political scientists' attempts to dislodge data from social media companies to study politics has laid the groundwork for contemporary challenges with the AI companies (several of which are, of course, the same companies). As we noted in our last book, the reliance on these powerful private companies for some of the most important information related to political phenomena places academics in very difficult positions. These companies are monopolies that control the information relating to basic facts about contemporary society.

> As dangerous as these information monopolies may be for purposes of economic competition, such dangers are compounded when only those who work for the firms and share in their corporate missions are able to gain social insights from the data they possess. . . . [T]he platforms control the information that most richly describes politics and society and therefore the data necessary to make sound judgments across virtually all major policy domains." (Persily and Tucker, 2020, p. 320)

As with social media, the relevant question is not whether the important data will be gathered, but whether the only people who will have access to the data will be employees tied to the profit maximizing mission of the firms controlling the frontier models.

Consequently, the three paths that researchers pursued with respect to social media remain the ones available today: work without the cooperation of the platforms to collect data; work with the cooperation of the platforms to collect data; or work with governments to compel the release of privacy protected data in secure environments or aggregated data released publicly. All three of these approaches have their shortcomings and raise concerns as to whether they can ultimately be successful.[7] Our advice, therefore, remains the same: Researchers need to pursue all of these

[6] Hall, A. B. [@ahall_research]. (2026a, January 3). Claude code for research [Post]. X. https://x.com/ahall_research/status/2007221974947508303

[7] For example, the EU Digital Services Act provides a pathway for researchers to demand access to social media data, but European researchers have only begun to take advantage of it and are facing bureaucratic hurdles. No similar program for data access for AI companies currently exists.

pathways simultaneously or else we, as a mass public, will risk being seriously under-informed about the possibly massive impact of AI on politics and society.

We do, however, want to highlight one important distinction between data access challenges for social media data and AI data. Yes, both are controlled by the whims of enormously powerful companies. But with social media, it has been possible, through a variety of different tactics, to observe how individuals are interacting with social media platforms: what they are positing, what posts are generating higher levels of engagement, and in many cases even to observe the networks that users are choosing to join. Auditing algorithms, around social media feeds, for example, has been much more challenging. Generative AI – and especially for the "closed weight" commercial models – has a similar level of opacity around the inner workings of the technology.

However, generative AI companies are in the business of selling users access to the models, so researchers are better positioned to audit the models through repeated interactions with them. There are, of course, usage limitations and pricing with which researchers need to contend; but, for now at least, there are very real opportunities for researchers to probe how these models respond to different queries. Conversely, though, it is much harder for researchers to observe how people are interacting with these new AI platforms, as compared to social media, as most interactions with generative AI (at least for now) are not in public fora but rather in private one-on-one interactions between users and the models. Thus, researchers will need to lean much more heavily into voluntary data donations from users, with all of the associated challenges for inferences, or they will need to rely on the platforms themselves to share user logs with those on the outside, with all of the associated concerns that entails.[8]

Overview of the Chapters

The volume begins with the substantive domains in which AI is reshaping political life, before concluding with two chapters that examine how AI is transforming the practice of political science itself – its methods and its pedagogy.[9]

The volume opens with Archon Fung, Bailey Flanigan, Florian Foos, and Charles Stewart's examination of "Artificial Intelligence and Democracy: Campaigns, Elections, Movements, and Deliberation." The authors assess how AI alters what different political actors can do across four critical domains. In campaigns, they distinguish between efficiency gains in back-office operations and the more transformative capacity for "listening at scale." In election administration, they note

[8] As early as 2023, one of us (Tucker) was publicly calling for AI companies to facilitate data donation for research in their chatbot products, but so far there has been no movement in this direction. See Sanderson, Z., & Tucker, J. A. (2023, November 1). *Beyond red-teaming: Facilitating user-based data donation to study generative AI*. Tech Policy Press. https://www.techpolicy.press/beyond-red-teaming-facilitating-user-based-data-donation-to-study-generative-ai/

[9] In an attempt to model the transparency around AI usage for which we are calling later in this chapter, we note that many of the chapter summary paragraphs below were drafted by AI (Claude) and edited by the chapter authors.

the potential for AI to relieve resource-constrained officials while cautioning against the risks of vendor dependence and hallucinated voting information. For social movements, the picture is mixed: AI lowers the barrier to producing high-quality content but also empowers activists to analyze government data and coordinate action. Most provocatively, the chapter argues that AI may solve the historic problem of scale in citizen deliberation. It points to experiments like Taiwan's "vTaiwan" initiative and Google DeepMind's "Habermas Machine" as evidence that AI-facilitated deliberation could move from niche experiment to central democratic practice.

In their chapter "Easy to Produce, Hard to Persuade: The Asymmetric Effects of AI on the Online Information Ecosystem," Brendan Nyhan, Jennifer Pan, Alexandra Siegel, and Yamil Velez provide a critical corrective to the prevailing panic about the persuasiveness of AI-generated content. Their central argument is one of asymmetric effects: While AI drastically lowers the cost of producing content, it does not solve the problem of persuading audiences. Most people consume relatively little political news and have strong partisan filters. The direct electoral impact of deepfakes and AI-generated text may, therefore, be less catastrophic than commonly assumed. The chapter instead identifies subtler dangers: the relaxation of constraints for low-capacity actors such as smaller states and extremist groups, the risk that state control over AI development embeds preferred narratives into the technology itself, the degradation of information environments in low-resource languages, and the "liar's dividend," whereby the mere existence of AI-generated content provides a ready-made excuse for dismissing authentic evidence. The chapter concludes by discussing AI's implications for research, both as a tool and as an object of inquiry in its own right.

What does it mean to study public opinion if the responses being studied reflect neither the "public" nor "opinion"? Joshua D. Clinton, Soubhik Barari, Ethan Busby, Trent D. Buskirk, Ray Duch, Anna-Carolina Haensch, D. Sunshine Hillygus, Courtney Kennedy, Kevin Munger, Doug Rivers, and Sean Westwood tackle these foundational challenges in "Public Opinion in the Age of AI." The chapter shows how the largest challenges in the use of AI to study public opinion are perhaps not technical, but conceptual. Although AI certainly portends the possibility of intriguing and important refinements - if not wholesale transformations – in how political scientists may choose to measure and characterize public opinion, the chapter also highlights the epistemological challenges that the use of some tools pose. Arguably the deepest risk is not technical errors, but the implication of substituting AI-generated responses for contemporaneous human voices. Given the foundational role that public opinion has for core notions of representation, accountability, and consent, it is profoundly important to raise questions about how the use of AI may affect the very notion of public opinion, in addition to its measurement. Transparency in the use of AI is of utmost importance when its use risks distorting our understanding of public opinion.

Baobao Zhang, Diane Coyle, Jae Yeon Kim, Johannes Himmelreich, and Mila Gascó-Hernandez shift the focus to the state itself, in "AI, the Public Sector, and Policymaking." The chapter presents a taxonomy of governmental AI uses, mapping thirteen categories of activity onto three core governance functions (policymaking, public service delivery, and internal management) with examples from governments worldwide. The authors identify three interconnected challenges: the

tension between public service values and algorithmic optimization, the growing information and power asymmetry between governments and private vendors rooted in decades of outsourcing that have hollowed out the state's technical capacity, and the resulting accountability gaps across vertical, horizontal, and diagonal dimensions of democratic governance. Drawing on cases ranging from Michigan's MiDAS unemployment system to Indiana's IBM welfare automation scandal and many others worldwide, they demonstrate what goes wrong when governments deploy AI without clear governance frameworks and strong state capacity. They argue that transparency, accountability, and enforceable constraints are central to a use of AI in the public sector that benefits citizens while reducing risks.

Rachel Gillum, Gregory Leslie, and Cara Wong examine "AI, Race, and Politics," situating artificial intelligence as both a reshaping force in political life and a system that reflects existing structures of racial and ethnic inequality. The chapter argues that bias in AI is often structural rather than incidental, emerging from models trained on data generated in stratified environments, even as AI can, under the right governance and design conditions, expand access and lower barriers to participation. Across domains of state governance, political communication, and electoral administration, the authors trace how AI systems redistribute visibility, voice, and administrative burden in ways that may reinforce or reconfigure racialized patterns of power. The chapter also demonstrates that AI is not only an object of political analysis but an increasingly central component of how political knowledge is produced, functioning as a measurement instrument whose errors are systematic and often correlated with the groups under study. In response, the authors introduce the AI Measurement Statement (AIMS), a four-question disclosure framework that makes visible how AI systems operate as measurement tools, where they introduce group-differentiated error, and how those risks shape inference. AIMS establishes a baseline expectation for how political scientists document and evaluate AI-mediated measurement and defines a standard for transparent, credible, and cumulative research across the discipline.

Dawn Teele, Shira Pindyck, and Sophia Lipkin explore "Artificial Intelligence and the Study of Gender and Politics," framing AI as a general-purpose technology that is reshaping the gendered distribution of power. The authors articulate a three-part framework: AI as a tool for data creation, a method of data analysis, and a category of analysis in its own right. As a tool, AI offers powerful new ways to reveal gender inequality – from computer vision analysis of non-verbal communication in debates to natural language processing that detects subtle biases in legislative text. As a category of analysis, the chapter asks how AI will reshape the material conditions of women's lives, raising the prospect of a new "breadwinner backlash" as automation transforms white-collar industries. The chapter also documents the rise of deepfake-enabled gender-based violence as a distinctly new form of political harm.

Carles Boix, Michael Becher, Valentina González-Rostani, and Daniel Stegmueller examine the political and institutional consequences of AI's economic impact in "AI's Economy and Its Political and Institutional Consequences." Unlike previous waves of information technology that primarily displaced routine manual labor, the authors argue, AI threatens to substitute for high-

skill, non-routine cognitive tasks – the work of the educated professional class. The chapter presents evidence that exposure measures for AI-related job displacement are surprisingly poorly correlated, yielding contradictory forecasts about which sectors are most vulnerable. The chapter then points out that AI's consequences for employment, wages, and collective bargaining are likely to be mediated by labor-market institutions and the pace of technological adoption. More consequential, the authors argue, is the concentration of capital. The training of frontier AI models requires immense computational resources, accelerating the concentration of economic power in a small number of firms, increasing business investment in fixed assets, and intensifying the incentives of big technology firms in regulatory capture. This fusion of capital concentration and potential labor displacement, they warn, may pose a substantial challenge to the social consensus on which democratic capitalism depends. It may also deepen global inequalities by encouraging reshoring and widening the technological gap between advanced and developing economies.

Sarah Kreps, Ben Buchanan, Michael Horowitz, and Erica Lonergan examine the international dimension of artificial intelligence in "AI, Geopolitics, and National Security." They frame AI as a system-altering, general-purpose technology that differs from earlier military innovations in both its diffusion and its integration across civilian and military domains. Unlike nuclear weapons, AI is dual-use, commercially embedded, and widely accessible, making it difficult to separate economic activity from military capability. The chapter argues that competition over AI increasingly turns on control over material and organizational foundations such as semiconductors and compute infrastructure rather than algorithmic breakthroughs alone. Yet rapid diffusion and the difficulty of measuring capabilities generate persistent uncertainty about relative advantage, intensifying security dilemma dynamics. In turn, because AI systems are opaque and resist direct verification, governance is likely to rely on indirect mechanisms such as export controls, standards, and coordination, rather than the formal arms-control regimes of the Cold War.

In "AI and Political Theory," Linda Eggert, Rob Reich, Ting-an Lin, Jeff Howard, and Lorenzo Manuali engage with the normative foundations of the discipline. The chapter argues that AI is forcing a convergence of theoretical traditions that have historically operated in isolation. Understanding algorithmic injustice, for example, requires blending insights from distributive justice, relational equality, and epistemic justice. The authors grapple with the "responsibility gap" that emerges when autonomous systems cause harm – can liability be traced to the programmer or does the autonomy of the system create a moral void? And they explore whether AI-facilitated deliberation can enhance democratic legitimacy or whether the mediation of the public sphere by non-human agents inevitably erodes the communicative action that democratic theory requires.

The volume's final two chapters turn to the discipline itself. In "AI and Research Methods," Christopher Barrie, Arthur Spirling, Alexis Palmer, Molly Roberts, James Bisbee, Jon Mellon, Lisa Argyle, Michael Heseltine, and Christopher Lucas provide a comprehensive survey of how large language models are being integrated into the political scientist toolkit. They identify six major applications – annotation and measurement, experimental treatment generation, silicon sampling, generative agent-based modeling, tool-augmented data collection, and the study of

LLMs as political objects in their own right – and assess the methodological challenges each presents. The chapter argues that LLMs do not simply make existing methods faster; they alter the standards by which validity, reproducibility, and transparency must be evaluated, and they recommend new practices of documentation and disclosure.

John Ishiyama, Christine Cahill, Jennifer De Maio, Stefan E. Kehlenbach, Sing Hui Lee, Steven Michels, Charles C. Turner, and Nicole Wu close the volume with a discussion of disruption of political science pedagogy in "Teaching and Learning Political Science in the Era of AI." Drawing on surveys of the APSA membership and a systematic review of the emerging literature, the authors document a discipline in transition: Most faculty view AI tools with skepticism but recognize the futility of prohibition. Instead, political science instructors should focus on teaching students the ethical and responsible use of AI. The chapter argues for a shift from product-oriented to process-oriented pedagogy, moving away from the "disposable assignment" toward scaffolded, project-based learning that makes the process of intellectual creation visible. Crucially, the authors frame AI literacy not merely as a technical skill but as an ethical and civic imperative, one that requires students to grapple with the environmental costs, labor practices, and power asymmetries embedded in these systems.

Cross-Cutting Themes

Although each of the chapters stands on its own and investigates very different dimensions of the relationship between AI, politics, and political science, we want to end this introduction by identifying some themes that cut across these different topics. The chapters identify several common dynamics relating to AI's effect on politics, as well as common challenges and opportunities that AI presents for political science. We see four themes, each with a dimension relevant to politics and research: (1) the political dynamics posed by the agglomeration of unprecedented power among a select few American corporations, (2) the opacity of AI models and difficulty replicating AI results, (3) inherent biases of AI models and systems, and (4) the need for appropriate benchmarks and standards.

The Unique Power of AI Corporations

The political challenge posed by powerful private actors in capitalist systems does not represent an unfamiliar phenomenon for political scientists. Indeed, depending on how broadly (or historically) one considers these issues, the study of corporate influence on politics stretches back to Karl Marx and Adam Smith, if not further. The impact of corporate power on politics is a recurring issue in research on campaign finance, lobbying, and the administrative state, let alone international relations scholarship on the role multinational corporations play in diplomacy, war, interstate inequality and development, and geopolitics generally.

All that said, the emerging dynamics posed by the power and influence of AI corporations represent an extreme difference in degree, if not in kind. To be sure, we have seen similar concerns raised regarding the immediately preceding generation of American tech companies, such as Meta, Google, Amazon, Apple, Twitter, and Microsoft, as they dominated various parts of the international technology stack and projected their power and values on topics as central to politics as permissible speech. The new dynamics, represented by many of these same companies, but also including OpenAI, Anthropic, Nvidia and others in the U.S. and China, present the spectre of companies with technology capable of influencing entire economies and political systems. Indeed, "influence" may understate the relationship, as these companies will be in the position of providing (or not) everything from the infrastructure to the technological capabilities for all governmental and economic sectors.

The dependence of governments on these corporations seems qualitatively different from classic examples (think Monsanto, Exxon, Dole, or Nestlé) in which critics raised concerns about foreign corporations holding less powerful governments hostage. As Anthropic's warnings about the cyberoffensive capabilities of its most recent models and its preceding conflict with the U.S. Department of War attest, these companies are developing technology with such wide ranging capabilities that they could literally bring down economies and political systems. Add to that the prospect of major economic shocks through transformations of the future of work and the labor market, and the politics of AI look quite different from the classic fears of corporate control of politics. With the United States federal government largely opting for a laissez faire approach to regulating these companies – even going so far as to threaten other countries that might adopt more aggressive policies – governments are now placed in a uniquely difficult position of deciding whether to accede to the wishes of these U.S. and Chinese companies or to be left behind in the AI economy with severe domestic political consequences.

The power of these companies is also relevant for political science research. As we noted above, these companies control access to data that will be relevant to contemporary politics and policy debates. This includes, of course, data concerning the use of AI in politically relevant settings and otherwise. But, for now, researchers also remain dependent on the companies to provide research-relevant information about, for example, the economic impact of AI, as exemplified by Anthropic's Economic Index, to name just one voluntary source of information. Researchers will then be either dependent on the good graces of companies to provide public information or they will run into conflict-of-interest concerns when they work directly and privately with them. Furthermore, even if AI companies choose to provide these kinds of "learned metrics" from people's usage of AI, the raw data of users' interactions with models – and models' responses to these real world interactions – is, absent meaningful regulatory interventions, likely to remain locked up inside these companies for reasons discussed previously. In sum, both governments and researchers are uniquely dependent on a small share of companies that dominate the AI economy.

The Opacity, Unpredictability, and Non-replicability of AI Model Outputs

The power of AI companies is compounded by the nature of the technology, the opacity of which presents unprecedented challenges related to interpreting how AI models perform. Here, too, we see impacts both for politics and for political science research. AI presents novel difficulties for governance of a fast-moving, opaque technology. It also poses unique challenges to the practice of political science, which requires replicability and interpretability.

With respect to governance and regulation, as suggested above, AI's unpredictability and opacity present challenges not seen with earlier technology. Nuclear technology, to which AI is often compared, was and is inspectable by governments. Whether for weapons or energy, officials inside and outside of government understand how the technology works and can regulate it accordingly. For AI, governments can test and red-team models, but auditing standards (as discussed more fully below) are necessarily incomplete with respect to the panoply of AI use cases (especially outside existential risks), slowly evolving, and not yet standardized. As a result, it seems literally impossible for evaluators – inside or outside government – to predict and assess model performance according to normal modes of governance applied to other technologies, like cars or planes. Short-staffed on relevant AI expertise, governments are, therefore, often dependent on representations made by model developers. Worse still, given the aggressive competition in the AI ecosystem, new models (or significant updates) are often released without warning and without adequate time to evaluate their performance or impacts. To be sure, some of these dynamics can be seen in other sectors of public administration, like drug development, financial products, or software; but, none of these settings present general purpose technologies with similarly far-reaching implications.

The challenges for political scientists are equally serious. Using AI for research can be fundamentally different from analyzing a large data set or conducting randomized controlled experiments. Generative AI models are, at their core, probabilistic models. This raises a question of how exactly do we think about the sensitivity or robustness of these models? Conceptually, we understand what we want to know here: How much do results change based on how we interact with these models? But probabilistic models should not be expected simply to produce the exact same output even when queried with the same prompt. There are some methods for making model outputs less variable (see, e.g., the discussion of "temperature" in Chapter 10), but some degree of variance in output is understood to lead to better performance. Thus generative AI introduces a trade-off between model performance and reproducibility, a characteristic not normally associated with classical conceptions of robustness testing. Depending on the nature of the study, moreover, maximizing consistency to ensure replicability may undermine the benefits of using AI in the first place.

Relatedly, how does one ensure that AI-augmented research is replicable if the generative models are (a) probabilistic and (b) changing all the time? We are encouraged by recent efforts by some of the authors in this volume to establish research practices and norms to deal with these

challenges, at least in part. "Guide-LLM" responds to some of the challenges above by attempting to establish standards of transparency surrounding LLM usage in research.[10] The categories of reporting the authors included in the checklist are:

- Scope of LLM Use
- Model/System Details
- Prompts
- Data Inputs and Privacy
- Validation and Interpretation
- Guidelines for Reproducibility
- Competing Interests

with an additional list of optional reporting:

- Justification for LLM Choice
- Rationale for Prompt Design
- Comparison Against Other Methods/LLMs
- Training Data Leakage Risks Addressed
- Risk of Bias or Systematic Differences Affecting Conclusions
- Conversation Transcripts
- Ethical Implications of the Research
- Computational Resources

Although this may seem cumbersome, we are encouraged by the ease with which agentic AI itself can likely be deployed to develop appropriate log files to allow for this sort of transparency. More generally, our hope is that, with this volume, we might begin to push the discipline into establishing these or similar norms for political science research employing AI, which may very well emerge as a substantial category of research in the field in short order.

Addressing Bias in AI Models and Related Research

Several of the chapters in this volume emphasize the dangers posed by bias in AI models. Unsurprisingly, the chapters relating to gender and racial politics do so extensively, but so do most of the other chapters in one or another respect. The chapters not only admonish users to be aware of the biases in these models and urge model developers to be transparent about these biases, but they also emphasize that bias in AI models is inevitable and baked into the process of training.

[10] Feuerriegel, S., Barrie, C., Crockett, M. J., Globig, L. K., McLoughlin, K. L., Mirea, D.-M., Spirling, A., Yang, D., Althoff, T., Antoniak, M., Argyle, L. P., Ashokkumar, A., Atari, M., Bailey, H., Bauer, K., Bhatt, U., Daumé, H., III, De Francisci Morales, G., Dehghani, M., ... Ribeiro, M. H. (2026). *A consensus-based reporting checklist for large language models in behavioral and social science*. https://llm-checklist.com/

Indeed, for both the politics of AI and related research, this volume represents a clarion call to recognize that bias is a feature of AI, not merely a bug.

The literature on AI bias is quite extensive, even if scholars ardently debate the appropriate ways to measure bias and which biases are politically consequential or worth addressing. (Wang et al., 2025) On the one hand, famous examples of AI's racial bias, such as the COMPAS recidivism algorithm deployed in the criminal justice system,[11] have highlighted the way that AI can be deployed in real world settings with racially discriminatory results. On the other hand, we have the notorious examples identified by critics of "woke AI," such as Google's early image generator tools that led to query responses that depicted black World War II German soldiers and a racially diverse set of Popes or Founding Fathers for the U.S. Constitution. Bias can creep into AI systems at every stage of their development – from the selection of data for training models to fine-tuning and to guardrails and filters that constrain an application's performance. Nevertheless, as the authors in the chapters that follow point out, a model that accurately represents the universe of content included in its training data will likely reflect the bias of the training data in its outputs.

Addressing AI bias is near the top of "political" considerations as policymakers within both government and the companies seek to erect guard rails against one or another bias. As AI tools roll out in arenas with well-developed antidiscrimination regimes, such as employment, housing, and education, existing laws need to be adapted to a world with human decisionmakers less "in the loop." Some jurisdictions, such as Colorado, which passed in 2024 the Colorado Anti-Discrimination in AI Law, have enacted broader laws that regulate AI as a technology to prevent discrimination against identified groups. Subsequently, the December 11, 2025, Executive Order, "Ensuring a National Policy Framework for Artificial Intelligence,"[12] specifically mentions the Colorado law, as an example of state laws that "are increasingly responsible for requiring entities to embed ideological bias within models." Thus, legal attempts to counteract bias are themselves accused of injecting other forms of bias – such is the politics of AI regulation in 2026 in the United States.

Of course, these concerns are not limited to the United States. The EU AI Act specifically requires high risk AI systems to examine "possible biases that are likely to affect the health and safety of persons, have a negative impact on fundamental rights or lead to discrimination."[13] At the other extreme, Chinese models, such as DeepSeek, appear to have been adjusted in the post-training phase of model development to ensure they refuse to answer (or give regime-favoring responses) to some politically sensitive questions.[14] However, even without these kinds of post-training

[11] Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). *Machine bias*. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[12] Executive Office of the President. (2025, December 11). *Eliminating state law obstruction of national artificial intelligence policy* [Presidential Action]. The White House. https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/

[13] EU AI Act, Article 10(2)(f).

[14] Li, J. (2025, January 31). What questions will China's DeepSeek not answer? *Deutsche Welle*. https://www.dw.com/en/what-questions-will-chinas-deepseek-not-answer/a-71470843

adjustments, if training data in the Chinese language tends to come disproportionately from government or government-approved sources, those political biases will likely "emerge" in model outputs, even in Western models. In addition to recognizing the underrepresentation of certain groups in training data that can lead to biased outputs, researchers need to be aware of the political realities that affect the very production of text in the first place that is used in model training. Social scientists, therefore, have an important role to play in understanding the contexts in which institutions affect the production of training data, and their concomitant consequences on the outputs of AI models.

Finally, beyond the question of bias itself lies the issue of *perceptions* of bias in the outputs of AI models. Much interesting research remains to be done on whether people tend to see AI generated news as inherently suspicious or, conversely, more trustworthy than traditional sources of news. Surely, there are important moderating factors to be considered here, such as the demographic characteristics or ideological proclivities of the individuals perceiving bias (or a lack thereof) in AI output. Perceptions of bias also may vary depending on the subject of the query and the related output. However, both for AI policy and for political science research, public perceptions of the bias of AI systems may be as important as the actual bias baked into the models or the training data.

Benchmarks and Evaluations

One of the challenges to addressing biases in AI models, as well as models' misalignment or poor performance, is the absence of agreed-upon metrics for evaluating AI systems. Any "problem" identified in performance assumes some judgment as to how a properly functioning system would behave and some method to detect when and how far a model strays from the ideal. The evaluation science of AI models has become a field unto itself, with potentially high stakes both for the development of models and the study of AI systems.

To some extent, the articulation of proper benchmarks is both the most minimal and most critical role that policy can play in shaping the future development of AI. Whether the area of concern is bioweapon development, the ubiquity of hallucinations, or the likelihood that a companion bot leads a vulnerable individual to commit suicide, judgments that a model erred assume some ground truth or agreement as to what proper performance entails. Even if we could agree on how a perfect model might perform, policymakers must decide how much error is tolerable, given the countervailing benefits of a given AI application. For example, autonomous vehicles will end up causing fatal accidents. But how much risk should be tolerated, especially given that human drivers may be equally or more dangerous on the roads? These decisions are inevitably political ones, with cost-benefit analyses that are not necessarily unique to AI. Yet anxiety about the rapid AI rollout and residual fears from the last decade's battles over technology and politics may lead some to expect more from AI systems than we do for their human counterparts.

Gaming of benchmarks has also now become a frequent critique lodged at any new model. Whether the metric of choice relates to general knowledge, mathematics, reasoning, coding, safety, or any number of other performance measures, once the "test" is well known, model developers can "teach to the test." The establishment of consistent benchmarks that can measure, over time, relative performance of different models has become both essential and bedeviling for anyone who cares about AI policy and governance. Indeed, for all the hyperventilation over the spectre of imminent development of AGI (Automated General Intelligence), agreement over what constitutes AGI has remained elusive.

We see the same kinds of problems in political science research. When evaluating the performance or propriety of AI tools for a given project, we must ask – as with the autonomous vehicle example above – "compared to what?" How do we know when an AI model or tool performs "well" for the researcher? For example, as referenced throughout this volume, one of the most important use cases of AI to date has been for classification tasks, such as categorizing text as about politics or not, reflecting an ideological position or not, being related to a particular topic area, etc. Instinctively, human coding holds strong appeal as the gold standard against which AI must be benchmarked. Yet we know human coding for classification tasks is fundamentally inconsistent – otherwise we would not report intercoder reliability, as all human coders would be generating the same classifications. So if humans are flawed at classification tasks, is it possible that machines could do a better job? We already know that machines can perform classification tasks at a speed and scale that far outstrips humans. However, if we recognize the inherent shortcomings of human performance on similar tasks, against what benchmark should we evaluate AI performance? And how might such benchmarks differ across objective (e.g. is there a flag in the picture?) versus subjective (e.g., does the image convey empathy?) classification tasks? If machines are simply probabilistic models, though, what does it even mean for a machine to be "better" than a human at classification tasks that do not have a simple ground truth, such as classifying content as liberal or conservative or threatening or not? And even if it is possible to answer this question, once we conclude that machines are "better" at classification tasks than humans, then how can we validate any further improvements in classification once we have lost human coding as the gold standard? And if we no longer have gold standards for validation of new methods, what does this ultimately mean about the reproducibility of research?

Versions of these questions permeate many of the chapters of this volume, but perhaps none more so than the chapter on AI and public opinion, which presents this dilemma at its starkest. As political scientists, we have enjoyed decades of methodological innovation based on the indisputable premise that simply asking a bunch of people their opinions is a fundamentally flawed way to measure public opinion. Statistical innovation after innovation has been aimed at figuring out how to better infer aggregate public opinion out of survey questions, but at the end of the day, even the best methods just do a better job of quantifying uncertainty. How then ought we to benchmark "silicon samples" of AI agents in an attempt to use these samples for measuring public opinion? If they match (flawed) human measures of public opinion? If they are subject to the same

sources of error (e.g., question ordering effects) as humans? If they get the macro-level estimates of public opinion "right" even while getting micro-level measures "wrong"? All of these questions, and many more, require not just objective measures of whether benchmarks are being met but also subjective judgments of what should be measured – to say nothing of the normative judgment of whether human public opinion can or should even be measured using non-human agents.

Conclusion

It is difficult to describe the anxiety and excitement surrounding the editing of a book on AI and politics in 2026. As with evaluating the significance of AI, in general, none of the historical analogs quite captures the moment. Is AI like computers or the internet, turbocharging human capabilities and economic innovation? Or is it like television or even the printing press, which revolutionized public communication and education, with destabilizing effects on existing power arrangements? Is it akin to nuclear technology, as it holds a promise of significant human benefits alongside potential catastrophe? Or is it more like electricity or the steam engine, a general purpose technology with multitudinous applications that spur on an industrial revolution with accompanying social dislocation? Might AI represent an even more fundamental change, like the creation of the alphabet or discovery of agriculture, which represented stages in the early evolution of the human condition? Each of these analogies offers some insight into a dimension of this new technology, but none can fully capture the potential impacts on existing institutions, let alone the study of them.

Even if, as all the authors in this volume admit, it is difficult to predict the exact trajectory of this technology, we need to be grappling now with its implications for governance and the study of politics. It may be beyond our capabilities to assess the full range of political implications of AI, but political scientists have an obligation, we think, to track its development and to evaluate how political actors incorporate AI into familiar dynamics that political scientists study. Indeed, we also think that the discipline has an important role to play in providing the analysis that can help steer technological development and its governance toward socially productive ends.

In this sense, the volume that follows represents a call for academics to join in the battle both to understand and to participate in the revolution that is underway. The challenges and opportunities this new technology poses for the pursuit of knowledge occur at a time when the very notion of human expertise and authority, across any number of dimensions, is being questioned. We cannot sit idly by as society, let alone our students, begin to outsource to machines the analysis of social phenomena, including politics. We hope that this book represents the beginning of the effort both to analyze and to harness this powerful technology to better understand the world of politics. As politics remains a central domain and apparatus for the allocation of power in *human* society, those of us who study the impact of AI on political power need to use our minds and voices, even if inevitably augmented by the technology itself, to make sense of the new world this technology is creating.

Epilogue

On January 2, 2026, or approximately six months into the process of writing the chapters for this volume, Stanford political scientist Andrew Hall posted on social media that "Claude Code and its ilk are coming for the study of politics like a freight train."[15] The following day he shared online an empirical political science paper (replicating and extending one of his own prior papers) where the data collection, data analysis, and writing of the paper was all done entirely by Claude Code – in a matter of hours.[16] In the ensuing months, the freight train has gained speed as some political scientists have begun transforming their research pipelines with AI agents, leading to ongoing speculation about what this development would mean not just for political science research but for the very academic institutions that have long sustained that research.[17] Not only are political scientists studying the impact of AI on the political world, but AI is simultaneously changing the way political scientists study those same phenomena.

It no longer seems debatable that agentic AI will turbo-charge some forms of political science research. It should also improve our ability to produce higher quality research thanks to the avalanche of review tools we are now able to throw at our work in progress. In addition to being able to use agentic AI to design research, draft pre-analysis plans, collect data, and conduct statistical analysis, we can use it to check the robustness of those analyses, draft literature reviews, and compile first drafts of entire papers. We also now have the ability to use AI agents to simultaneously check our code for errors, check our writing for style, accuracy, and within-paper consistency, and check the validity of our causal claims, the fidelity of our analysis to pre-analysis plans, our bibliography references, and much more. And while not every recommended change to a paper suggested by an AI agent is going to be a good one, the sheer quantity of feedback we can now receive on our work in mere hours (minutes) is stunning.[18]

If we accept the assumption that AI can accelerate the production of high quality research, though, then it raises an important set of questions for political science as a discipline – and political science as a set of academic institutions designed to support research. We note in particular the following five issues:

[15]Hall, A. B. [@ahall_research]. (2026a, January 3). Claude code for research [Post]. X. https://x.com/ahall_research/status/2007221974947508303

[16] Hall, A. B. [@ahall_research]. (2026b, January 4). Proof of Claude Code for empirical polisci paper [Post]. X. https://x.com/ahall_research/status/2007603340939800664

[17] Messing, S., & Tucker, J. A. (2026, March 3). *The train has left the station: Agentic AI and the future of social science research*. Brookings Institution. https://www.brookings.edu/articles/the-train-has-left-the-station-agentic-ai-and-the-future-of-social-science-research

[18] For just one excellent example, see Bäckman, C. (2025). *AI-research-feedback*. GitHub. https://github.com/claesbackman/AI-research-feedback

1. What happens to our journals and our existing system of peer review as the time necessary to complete a sufficiently high quality paper to submit to a refereed journal drops from months or years to weeks or days?
2. How might scholars *consume* new research, already a challenging task, if the speed at which it is produced continues to accelerate?
3. How might training of students evolve when so much of the work for which we previously employed research assistants – both undergraduate and graduate – can now be done much faster and cheaper by agentic AI?
4. Will AI tools that accelerate some forms of research more than others ultimately end up privileging the former over the latter? Might decisions about hiring and promotion end up privileging the type of work that can be accelerated by AI, or, conversely, doing the opposite by rewarding research that can not be replaced by AI (e.g., field work)?
5. More generally, how might hiring and promotion standards react to rapid increases in productivity, and how will existing methods of evaluation of scholarly output evolve to address these productivity gains?

Taken together, it is possible to imagine a future where largely AI produced papers are reviewed by AI agents, and then, once posted or published, are read and summarized by other AI agents for academics to peruse before being passed to yet another set of AI agents to generate ideas for new research papers as the cycle begins anew. Where exactly will we, as human political scientists, ultimately fit into this process?

We have not yet reached – and indeed may never reach – this version of a future for our profession where humans are increasingly crowded out of the research process by AI agents. As with its effect on all forms of labor, AI might be equally likely to generate a virtuous cycle in which lowering the cost of research leads to increased production of political science, which, in turn, leads to greater demand for this kind of work – a sort of "Jevon's paradox" for political science. Or perhaps advocates of the view of AI as a "normal technology" are correct in this realm, as well, and productivity gains may only be realized far in the future, if at all, as it takes a generation for this technology to diffuse throughout the profession. But we are at the point where we at least need to begin thinking about these questions. In the meantime, as this volume richly demonstrates, there is much research to be done about the role of AI in politics and political science.

*****

AI Disclosure Statement for this chapter: AI was used to draft the paragraph summaries of the chapters found in the "Overview of the Chapters" section. These summaries were then edited by the subcommittee chairs as well as by Persily and Tucker. All errors remain our own.

References

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). *Machine bias*. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (Accessed: [April 21, 2026]).

Bäckman, C. (2025). *AI-research-feedback* [Computer software]. GitHub. https://github.com/claesbackman/AI-research-feedback (Accessed: [April 21, 2026]).

Capstick, E. (2025). Public opinion. In *AI index report 2025* (Chap. 8). Stanford Institute for Human-Centered Artificial Intelligence. https://hai.stanford.edu/assets/files/hai_ai-index-report-2025_chapter8_final.pdf (Accessed: [April 21, 2026]).

Diamond, L. (2010). Liberation technology. *Journal of Democracy*, *21*(3), 69–83.

European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council (EU AI Act). *Official Journal of the European Union*, *L 2024/1689*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (Accessed: [April 21, 2026]).

Executive Office of the President. (2025, December 11). *Eliminating state law obstruction of national artificial intelligence policy* [Presidential Action]. The White House. https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/ (Accessed: [April 21, 2026]).

Feuerriegel, S., Barrie, C., Crockett, M. J., Globig, L. K., McLoughlin, K. L., Mirea, D.-M., Spirling, A., Yang, D., Althoff, T., Antoniak, M., Argyle, L. P., Ashokkumar, A., Atari, M., Bailey, H., Bauer, K., Bhatt, U., Daumé, H., III, De Francisci Morales, G., Dehghani, M., ... Ribeiro, M. H. (2026). *A consensus-based reporting checklist for large language models in behavioral and social science*. https://llm-checklist.com/ (Accessed: [April 21, 2026]).

Hall, A. B. [@ahall_research]. (2026a, January 3). Claude code for research [Post]. X. https://x.com/ahall_research/status/2007221974947508303 (Accessed: [April 21, 2026]).

Hall, A. B. [@ahall_research]. (2026b, January 4). Proof of Claude Code for empirical polisci paper [Post]. X. https://x.com/ahall_research/status/2007603340939800664 (Accessed: [April 21, 2026]).

Li, J. (2025, January 31). What questions will China's DeepSeek not answer? *Deutsche Welle*. https://www.dw.com/en/what-questions-will-chinas-deepseek-not-answer/a-71470843 (Accessed: [April 21, 2026]).

Messing, S., & Tucker, J. A. (2026, March 3). *The train has left the station: Agentic AI and the future of social science research*. Brookings Institution. https://www.brookings.edu/articles/the-

train-has-left-the-station-agentic-ai-and-the-future-of-social-science-research/ (Accessed: [April 21, 2026]).

Munger, K., Bakker, B. N., Berinsky, A. J., Just, N., Guess, A. M., Giger, N., Tenenboim-Weinblatt, K., Lawrence, R., & van de Rijt, A. (2026). *Peer review 2027: Scenarios for academic publishing in the age of AI* [Preprint]. SocArXiv. https://osf.io/preprints/socarxiv/594zj_v1 (Accessed: [April 21, 2026]).

Munger, K. (2023). Temporal validity as meta-science. *Research & Politics*, *10*(3). https://doi.org/10.1177/20531680231187271 (Accessed: [April 21, 2026]).

Persily, N., & Tucker, J. A. (Eds.). (2020). *Social media and democracy: The state of the field, prospects for reform*. Cambridge University Press.

Quinnipiac University. (2026, March 30). *The age of artificial intelligence: Americans' AI use increases while views on it sour* [Poll]. https://poll.qu.edu/images/polling/us/us03302026_uaio53.pdf (Accessed: [April 21, 2026]).

Sanderson, Z., & Tucker, J. A. (2023, November 1). Beyond red-teaming: Facilitating user-based data donation to study generative AI. *Tech Policy Press*. https://www.techpolicy.press/beyond-red-teaming-facilitating-user-based-data-donation-to-study-generative-ai/ (Accessed: [April 21, 2026]).

Wang, A., Phan, M., Ho, D., & Koyejo, S. (2025). *Fairness through difference awareness: Measuring desired group discrimination in LLMs*. arXiv. https://arxiv.org/pdf/2502.01926 (Accessed: [April 21, 2026]).

Wike, R., Silver, L., Fetterolf, J., Huang, C., Austin, S., Clancy, L., & Gubbala, S. (2022, December 6). *Views of social media and its impacts on society in advanced economies*. Pew Research Center. https://www.pewresearch.org/global/2022/12/06/views-of-social-media-and-its-impacts-on-society-in-advanced-economies-2022/ (Accessed: [April 21, 2026]).