

Executive Summary

Artificial intelligence (AI) promises to transform how government agencies do their work. Rapid developments in AI have the potential to reduce the cost of core governance functions, improve the quality of decisions, and unleash the power of administrative data, thereby making government performance more efficient and effective. Agencies that use AI to realize these gains will also confront important questions about the proper design of algorithms and user interfaces, the respective scope of human and machine decision-making, the boundaries between public actions and private contracting, their own capacity to learn over time using AI, and whether the use of AI is even permitted. These are important issues for public debate and academic inquiry.

Yet little is known about how agencies are currently using AI systems beyond a few headline-grabbing examples or surface-level descriptions. Moreover, even amidst growing public and scholarly discussion about how society might regulate government use of AI, little attention has been devoted to how agencies acquire such tools in the first place or oversee their use.

In an effort to fill these gaps, the Administrative Conference of the United States (ACUS) commissioned this report from researchers at Stanford University and New York University. The research team included a diverse set of lawyers, law students, computer scientists, and social scientists with the capacity to analyze these cutting-edge issues from technical, legal, and policy angles. The resulting report offers three cuts at federal agency use of AI:

- a rigorous canvass of AI use at the 142 most significant federal departments, agencies, and sub-agencies (Part I)
- a series of in-depth but accessible case studies of specific AI applications at seven leading agencies covering a range of governance tasks (Part II); and
- a set of cross-cutting analyses of the institutional, legal, and policy challenges raised by agency use of AI (Part III).

Taken together, these analyses yield five main findings.

First, the government's AI toolkit is diverse and spans the federal administrative state. Nearly half of the federal agencies studied (45%) have experimented with AI and related machine learning (ML) tools. Moreover, AI tools are already improving agency operations across the full range of governance tasks, including:

- *Enforcing* regulatory mandates centered on market efficiency, workplace safety, health care, and environmental protection;

- *Adjudicating* government benefits and privileges, from disability benefits to intellectual property rights;
- *Monitoring and analyzing* risks to public health and safety;
- *Extracting* useable information from the government's massive data streams, from consumer complaints to weather patterns; and
- *Communicating* with the public about its rights and obligations as welfare beneficiaries, taxpayers, asylum seekers, and business owners.

The government's AI toolkit spans the full technical scope of AI techniques, from conventional machine learning to more advanced "deep learning" with natural language and image data.

The government's AI toolkit is diverse and spans the federal administrative state. Nearly half of the federal agencies studied (45%) have experimented with AI and related machine learning (ML) tools.

Second, and despite wide agency embrace of AI, the government still has a long way to go. In canvassing agency use of AI, Stanford computer scientists evaluated the techniques deployed in each use case and, while limited public details precluded clear conclusions as to many, rated only 12% as high in sophistication. This is concerning because agencies will find it harder to realize gains in accuracy and efficiency with less sophisticated tools. This result also underscores AI's potential to widen, not narrow, the public-private technology gap.

Third, AI poses deep accountability challenges. When public officials deny benefits or make decisions affecting the public's rights, the law generally requires them to explain why. Yet many of the more advanced AI tools are not, by their structure, fully explainable. A crucial question will be how to subject such tools to *meaningful accountability* and thus ensure their fidelity to legal norms of transparency, reason-giving, and non-discrimination. The case studies presented in the report highlight several vital aspects of that challenge:

- Transparency's costs, benefits, and feasibility will vary across policy areas, governance tasks, and AI techniques. Open-sourcing of technical details might be appropriate when agencies are allocating social welfare benefits but can undermine agency use of valuable enforcement tools because of gaming by regulatory targets.
- One key area for future inquiry is how to adapt existing principles of administrative law, which is more likely to modulate agency use of AI than the constitutional constraints that occupy much current debate.
- Policymakers should also consider other interventions. A promising candidate is to require agencies to engage in prospective "benchmarking" of AI tools by reserving a random hold-out sample of cases for human decision, thus providing critical information to smoke out when an algorithm has gone astray or "automation bias" has led decision-makers to excessively defer to an algorithm.

To achieve meaningful accountability, concrete and technically-informed thinking within and across contexts—not facile calls for prohibition, nor blind faith in innovation—is urgently needed.

Fourth, if we expect agencies to make responsible and smart use of AI, technical capacity must come from within. While many agencies rely on private contractors to build out AI capacity, a majority of profiled use cases (53%) are the product of in-house efforts by agency technologists. This underscores the critical importance of internal agency capacity building as AI continues to proliferate. In particular:

- In-house expertise promotes AI tools that are better tailored to complex governance tasks and more likely to be designed and implemented in lawful, policy-compliant, and accountable ways. Sustained collaboration between agency officials and in-house technologists facilitates identification of appropriate questions, seizing new innovations, and evaluating existing tools, including contractor-provided ones.
- Fully leveraging agency use of AI will require significant public investment to draw needed human capital and update outmoded data and computing systems. Given fiscal and labor market constraints, agencies should also explore non-commercial sources of valuable technical capacity, including collaborations with universities, NGOs, and industry and agency-sponsored competitions.

In-house expertise yields AI tools that are better tailored to complex governance tasks and more likely to be implemented in a lawful, policy-compliant, and accountable fashion.

Fifth, AI has the potential to raise distributive concerns and fuel political anxieties. Growing agency use of AI creates a risk that AI systems will be gamed by better-heeled groups with resources and know-how. An enforcement agency's algorithmic predictions, for example, may fall more heavily on smaller businesses that, unlike larger firms, lack a stable of computer scientists who can reverse-engineer the agency's model and keep out of its cross-hairs. If citizens come to believe that AI systems are rigged, political support for a more effective and tech-savvy government will evaporate quickly.

To achieve meaningful accountability, concrete and technically-informed thinking within and across contexts—not facile calls for prohibition, nor blind faith in innovation—is urgently needed.

In sum, the stakes are high. Managed well, algorithmic governance tools can modernize public administration, promoting more efficient, accurate, and equitable forms of state action. Managed poorly, government deployment of AI tools can hollow out the human expertise inside agencies with few compensating gains, widen the public-private technology gap, increase undesirable opacity in public decision-making, and heighten concerns about arbitrary government action and power. Given these stakes, agency administrators, judges, technologists, legislators, and academics should think carefully about how to spur government innovation involving the appropriate use of AI tools while ensuring accountability in their acquisition and use. This report seeks to stimulate that thinking.