

# Democracy and the AI-Augmented Administrative State

Johannes Himmelreich

## Introduction

Stanislav Petrov's name stands for an idea. He is the man who, on September 26, 1983, saved the world. As alarms rang out that five U.S. nuclear missiles were on their way to Russia, Petrov, who was tasked with escalating this alarm and initiating a counterstrike, decided to wait. He considered what was at stake and suspected the alarm was a false one.

His name today stands for the idea of integrity, and specifically of doing a *good job*. Arguably, however, Petrov did a good job by *not* doing what he was supposed to do; that is, he *did not* do well on his assigned *task*.

Petrov's task that night in 1983 was to analyze signals in incoming radar data, discern patterns, and take actions in response. This is the kind of task that, today, artificial intelligence (AI) does better than humans. Wherever humans perform such data-analysis tasks today, they stand to be augmented or replaced by machines tomorrow.

Augmenting or replacing public servants like Petrov carries a risk that is widely overlooked. Even if Petrov's professional descendants today may not be as good at their *task* as machines, they might still be very good at their *job*.

AI hence presents a challenge for how we think about administrative capacity. The distinction between *performing a task well* and *doing a good job*, whatever it is, is not typically drawn when administrative capacity is evaluated. Insofar as this distinction matters, however, overlooking it carries the risk of AI inadvertently diminishing the quality of public service delivery.

Without a firm grasp on the difference between merely performing a task and doing a good job, we might be unable to articulate an important aspect of administrative capacity. When AI is introduced into public administration absent a clear view of what matters—for doing a good *job*—AI will optimize an impoverished objective. AI will yield measurable improvements regarding how well *tasks* are performed while, at the same time and unseen, undermining the very values involved in the broader idea of doing a good *job*. AI hence threatens effective administrative capacity. In this chapter, I investigate this distinction between task and job.

The distinction between seeing public servants merely as performing tasks vs. doing a good job makes a difference for democratic legitimacy. Public servants who do a good job safeguard stability, innovate rulemaking, and resist political pressure. For example, when public servants stay on the job across governments with opposing agendas, they ensure that rules and regulations are applied consistently. They thus uphold an element of the rule of law. Public servants can exercise bureaucratic discretion in a way that contributes to democratic legitimacy. Or, for another example, career civil servants do not need to raise funds for their reelection; many civil servants enjoy job security that renders them immune against retaliation from the administrations that they serve. This gives them greater independence regarding antitrust regulation, such as approving mergers.

This chapter investigates how public servants exercise bureaucratic discretion to uphold democratic values—public servants who “do a good job.” This is an overlooked aspect in the evaluation of administrative capacity. After illustrating the difference between a task and a good job, I theorize this difference in terms of what I call “norms of responsible public service.” I suggest that these norms provide a standard by which to evaluate how governments should use AI.

I concentrate on the public sector, particularly the administrative state or the bureaucracy. This is where the stakes are often high and, moreover, the difference between *performing a task well* and *doing a good job* is somewhat easier to see. Still, the distinction generalizes. The risk that AI adoption erodes professional norms while improving task performance exists *mutatis mutandis* in the private sector as well.

Moreover, I concentrate on the risks AI poses to democracy, via the administrative state, as opposed to its benefits, thus fitting with the theme of this volume. Arguably, AI offers risks and benefits for democracy (Kreps and Kriner 2023). I discuss its potential benefits elsewhere (Himmelreich 2023). Today, the risks have received a fair hearing. Broadly, AI threatens democracy because AI can undermine democratic deliberation (Allen and Weyl 2024), facilitate oppression, surveillance, and human rights abuses, or because it might cause an existential catastrophe (Bengio 2023). However, the perspective of this chapter, that AI threatens democracy by undermining features of the administrative state that are important and valuable, is an aspect that has received less attention than it deserves (Bullock 2019).

## AI in Government

The administrative state, or the bureaucracy, consists of the executive branch of government, which includes career civil servants in different departments, independent agencies, and offices at different levels. More broadly, it includes teachers, military personnel, advisors, political

appointees, and elected officials. Analysts have long argued that AI “has the potential to enhance almost everything government does” (Eggers and Beyer 2019). Since then, the U.S. has come to lead the global index of government AI readiness by investing aggressively in the governance, innovation, and larger ecosystem of AI (Hankins 2023).

The “visible” federal government (that is, without functions related to national security) lists 710 AI use cases in its AI inventory (Federal Chief Information Officers Council 2023). Not all of these uses of AI are directly linked to decision making. Among the 710 use cases are a model to extract networks of sidewalks from street-level images like Google Street View (used by the Centers for Disease Control and Prevention), a model to detect whether someone is infected with Covid-19 or the influenza virus just by the sound of their cough, and a chat bot, Aidan, used by the Department of Education to answer common questions about financial aid.

The distinction between task and job might be relevant to such mundane uses of AI in government, but the cases I have in mind—the paradigmatic cases—are those where governments use AI to decide a course of action or to allocate scarce services. These are the kinds of cases I’ll mainly refer to with “AI in government” or “government AI.”

Since 2023, the federal government has been using AI widely to make administrative decisions. For example, the Social Security Administration (SSA) uses AI to predict which disability claims are going to be accepted, the US Citizenship and Immigration Services (USCIS) to identify text patterns that indicate fraud, and local governments which streets need to be repaired (Engstrom et al. 2020; Glaze et al. 2024; Federal Chief Information Officers Council 2023; DHS 2025). Moreover, government AI is used in enforcement decisions (e.g., pre-trial risk scoring) and management decisions (e.g., using natural language processing to analyze and revise existing regulations) (Engstrom et al. 2020).

Enthusiasm about AI as a tool for reforming the bureaucracy has increased markedly in the second Trump administration. In contrast, transparency about where the federal bureaucracy uses AI has decreased. Likewise, public information about how AI is used by the armed forces or at other levels of government is typically very limited. Technically, it would be feasible today to use AI, at the state level, for example, to adjudicate claims for unemployment insurance. It could also identify families for special services like youth intervention programs, and within school districts, to help grading, identify potential dropouts, or nominate students for gifted-and-talented programs. However, the extent to which AI *is* in fact used for these or other tasks is unclear. This lack of transparency creates a concerning deficit of democratic legitimacy.

## “Good Job”—Responsible Public Service

Consider the responsibilities of teachers, soldiers, and civil servants. Think of the federal workers at the Social Security Administration who adjudicate disability benefit claims. Or consider local officials, like mayors or employees in state taxation department. For each of these roles, you likely have an idea of what it would look like for them to do a good job. The claims adjudicator, for example, who reads up on rare medical conditions and experiences of disabled also outside of their work. Or a teacher who discreetly investigates the reasons behind a student’s sudden drop in performance. When someone in these roles does a good job, they are upholding what I call the norms of responsible public service. I suggest that *one* part of these norms of responsible public service plays an important role within a democracy. However, I worry that we don’t often clearly discern what responsible public service is and what it does. The measures with which we evaluate administrative capacity are incomplete. They fail to capture how the exercise of bureaucratic discretion can be a democratic practice. To illuminate this, let’s consider three cases that, in my view, are uncontroversial cases of responsible public service. To start, let’s return to Stanislav Petrov.

### Stanislav Petrov’s Courage

Petrov decided not to escalate the alarm that a nuclear strike was underway on that important night in September 1983. He reportedly reached this conclusion after considering all of the evidence available to him. He reasoned that although the radar system, named “Okó,” indicated an incoming first strike, what it showed didn’t make much sense. First, Okó initially detected the launch of only one missile (although another four thereafter), which is not what one would reasonably expect how a first strike would begin. Rather, it should have been an all-out attack. Consequently, Petrov reasoned that Okó must be malfunctioning. Second, Okó had been operationalized only relatively recently. Knowing this, Petrov didn’t fully trust it. Finally, Petrov considered an ethical question. He pondered whether he wanted to be responsible for “unleashing the third World War ... [and] I said, no, I wouldn’t.”<sup>1</sup>

This episode illustrates several aspects of what it means for a public servant to do a “good job.” Petrov’s story is usually told as one of courage—supposedly the courage to overrule the machine. As I see it, courage was indeed required as an enabling condition but, more fundamentally, this episode illustrates the exercise of expertise in interpreting evidence and reasoning in pursuit of stability and peace. Essentially, because his skepticism was fueled

---

<sup>1</sup> This incident is meant to serve as an illustration not evidence. It would be rather weak evidence (since it relies largely on Petrov’s recollection from memory and self-explanations after the fact) from sources with unclear provenance.

from various sources, Petrov held firmly onto certain previously held beliefs that there was not a credible first strike.

In short, Petrov concluded that his *own* assessment of the evidence and reasoning to be superior to that of Oko. Indeed, Petrov did have access to a broader range of evidence than Oko in three ways. First, Oko did *not* know that it was put into operation only recently. Petrov took such metadata into account. Second, Petrov had background knowledge of what an attack would look like. Oko did not. Oko presented incoming objects, initially one, then five. But Petrov could rely on his reason to deduce that what looked like incoming objects wasn't, in fact, conclusive

evidence of an attack. Finally, Petrov engaged in *moral* reasoning, namely, that he would not personally bear the responsibility for unleashing a nuclear war (especially not on the basis of the evidence that Oko offered).

Petrov's story hence highlights how public servants, when making administrative decisions, integrate a broad range of evidence (Oko's track record), add expertise by interpreting data (what an attack would look like), and supplement moral reasoning (considering the stakes of decisions). This is the difference between merely performing a task and doing a job and our first example of responsible public service. It illustrates an important feature of administrative capacity: that bureaucratic discretion can be exercised in a way that furthers important ends.

### The EPA's Radon Regulation

Compared to the case of Petrov, almost all examples of responsible public service are more mundane. Moreover, responsible public service is often a collective effort. Even when individuals make an outsized contribution, they typically are not known by name. Although awards for good public service *do* exist, as one observer notes, they are "awards that most people will never know were handed out" and offered by an "organization that no one's ever heard of" (Lewis et al. 2025). In one way, the awards for public service hence resemble good public service itself: public administration is, in practice, largely invisible but nevertheless the basis for "the welfare, happiness, and very lives of all of us" (Waldo 1955). As dramatic as this may sound, it is not an exaggeration.

A case that is as mundane as it is impactful is that of the regulation of radon emissions. Radon is an odorless and colorless gas that occurs naturally and is radioactive (a byproduct of decaying uranium). Radon has been known as an obscure hazard for miners since the 1500s, until, in December 1984, a construction engineer at a nuclear power plant in Pennsylvania triggered radiation alarms. He was neither working in a mine nor was he exposed to nuclear fuel. It turned out that the level of radon gas in his home was about 1,000-times higher than what eventually became the safe standard, or equivalent to smoking more than 100 packs of

cigarettes per day. Until then, it had not been considered possible that a radioactive gas could naturally occur at such a concentration within a building.

Without explicit congressional authorization, and using only its existing statutory authority (on “indoor air quality”), the Environmental Protection Agency (EPA) took up the issue. By August 1986, the EPA had coordinated a nationwide survey program for radon, reviewed the existing literature, established a recommended “action level” for indoor air, and published “A Citizen’s Guide to Radon” and “A Homeowner’s Guide to Radon”—all within little more than a year after the Pennsylvania incident.<sup>2</sup> Congress expanded the EPA’s statutory responsibility to explicitly include radon and authorized funding for the already on-going policy and research only in October 1986 with the Radon Gas and Indoor Air Quality Research Act.<sup>3</sup> In 1988, with the Indoor Radon Abatement Act, Congress eventually placed the existing EPA’s programs on a statutory basis and charged other governmental agencies to act. The story of how radon came to be regulated—driven by a bureaucratic initiative—played out similarly elsewhere, including in Canada (Harrison and Hoberg 1991; Heath 2020).

I take such regulation that arises from an independently pursued bureaucratic initiative to be a second example of public administrators doing a good job. Strictly speaking, radon wasn’t part of the EPA’s set of tasks, which included undertaking authorized programs and enforcing statutes about clean water, clean air, resource conservation, and the control of toxic substances.

Yet, even if what the civil servants at the EPA did wasn’t part of their established responsibilities, they arguably did a good job. For one, the marginally statutory initiative by civil servants in the EPA was essential. It is hard to see how the danger posed by radon otherwise would have found its way onto the political agenda, be investigated, and addressed so quickly if it were not the EPA’s bureaucratic initiative (Heath 2020, 37). Radon gas is not an issue that wins votes, differentiates candidates from their opponents, or offers politicians photo opportunities. Instead, the issue of radon is relatively technical and boring. As such, bureaucrats were well positioned to work on it.<sup>4</sup>

Moreover, the process by which radon gas came to be regulated highlights several norms of responsible public service. The bureaucrat-driven regulatory effort required expertise,

---

<sup>2</sup> The EPA’s archive of publications can be found here: <https://nepis.epa.gov>

<sup>3</sup> In its report to Congress in 1989, the EPA states that “the Radon Gas and Indoor Air Quality Research Act, for the first time gave EPA clear authority to begin to address indoor air quality problems on a more comprehensive basis.”

<sup>4</sup> Nevertheless, regulating radon is not an apolitical matter: regulating radon increases the administrative burden, the complexity of building codes and workplace safety regulations, and has significant compliance and enforcement costs.

evidence, a commitment to the public interest, and a long-term perspective. The EPA knew about radon, why it occurs, and how it impacts human health—although, for residential homes, the EPA had concentrated on exposure through tap water. Once the Agency became aware that radon can occur in high concentrations in residential buildings, the EPA used its expertise and connections to the academic community to conduct studies to gather better evidence on radon emissions in residential homes. Finally, the EPA employees demonstrated a commitment to long-term planning and advancing public interest (since, like smoking, the effects of radon exposure compound over long periods of time).

This story of radon gas generalizes not only to other countries but also to other policy domains. The celebrated role of the U.S. Federal Drug Administration in drug safety testing or the reform of Canadian railway subsidies are similar examples of civil servants, and not voters or legislators, making policy. In these and many other cases, bureaucrats pursued their own regulatory agenda *independently*, that is, following their *own* judgement of what, in the public interest, *should* be done (Heath 2020, chap. 1). In doing so, bureaucrats may subvert the policies of legislators or appointed officials in the executive (O’Leary 2019). Insofar as bureaucrats should only implement but not make policy, such bureaucratic initiative and independence appears undemocratic.

Notwithstanding concerns about democratic legitimacy, such bureaucratic initiative and independence can be democratic. Democracy has extra-populist elements, such as the protection of civil liberties, the integrity of the electoral process, a healthy political culture, a free press, rule of law, and a civic education.

Finally, and most pertinent for the purposes here, democracy also requires a sufficient level of *government functioning*. When governments fail to maintain infrastructure, enforce regulations, deliver services, or prevent unrest, democratic institutions lose legitimacy. When government is perceived as incompetent, unstable, or corrupt, citizens disengage or are attracted to authoritarianism or clientelism. As such, public servants who implement policy, maintain standards, ensure peace and stability, and solve collective action problems are part of the infrastructure that makes democratic governance possible. How the EPA came to regulate radon gas in residential buildings hence is an example of high administrative capacity—an exercise of bureaucratic discretion (at the organizational level) as a democratic practice.

This bureaucratic ingredient of democratic practices is recognized by democracy indices that measure how democratic different countries are. The Economist Intelligence Unit, for example, classifies both the U.S. and Singapore as equally flawed democracies. Although Singapore is lacking in citizen political participation, in the quality of its electoral process, and in the protection of civil liberties, it enjoys a much higher level of government functioning. (Singapore is still ranked as less democratic overall).

## General Milley's Assurances

Before delineating more systematically the various ways in which public administrators exercise bureaucratic discretion and contribute to administrative capacity by upholding the norms of responsible public service, consider a third and last case.

In the last weeks of the first Trump administration and the days following the storming on the U.S. capitol on January 6<sup>th</sup> in 2021, General Mark Milley, then the chairman of the Joint Chiefs of Staff, contacted his counterparts in China to assure them that the U.S. was not planning to conduct “kinetic operations.” During calls in October 2020 and January 2021, he reportedly said that “[i]f we’re going to attack, I’m going to call you ahead of time. It’s not going to be a surprise” (Goldberg 2023). To back up these assurances to China, Milley had planned to place himself into the chain of military command, where he—whose role, despite the uniform, is technically a civilian one as an advisor—has no formal place. By law, the chairman has no command authority. Providing an adversary with advanced notice of military operations against them would be treason. This was consistent with a certain behavioral pattern General Milley had displayed and the (dis-)reputation he had subsequently earned throughout the first Trump administration as, depending on who you ask, either a “hero protecting the constitution” or the “kraken of the swamp.”<sup>5</sup>

Like Petrov, Milley didn’t do his task well, but, I argue, he did do a good job. Milley pre-empted an escalatory cycle at a time when such a cycle seemed to be a serious possibility. By January 2021, several people in the administration expressed concern over the president’s impulsive and erratic leadership and questioned his fitness for office. Now the president seemed to endorse, if not encourage, a coup to prevent the certification of an election, which, at the time, was expected to end his political career. Throughout his tenure, the president had threatened North Korea with a nuclear attack (at one point boasting to “have a Nuclear button ... much bigger and more powerful” than Kim Jong Un’s). Recently, China reportedly received intelligence warning of a planned nuclear attack. After the attempted coup, domestic political leaders were also seeking reassurance that the U.S. nuclear arsenal was “stable.” Calls with military counterparts are a common practice and Milley claims to have coordinated these calls with the respective secretary of defense (Lubold 2021).

In sum, Milley’s actions are an example of responsible public service. Ensuring stability serves the public interest. Pre-empting an escalatory cycle prevents military confrontations, large and small, that threaten the lives of Americans—in active service or in the general population. Mistaken intelligence in an uncertain strategic environment can lead to a large-

---

<sup>5</sup> Liberal journalists (of The Atlantic) in the former, and Sebastian Gorka, Deputy Assistant to President Trump, in the latter instance (Goldberg 2023).

scale conflict today as much as when Stanislav Petrov didn't raise the alarm. Making assurances about strategic and political stability decreases the risk of an accidental escalation. Making such assurances credibly required a certain reputation and influence that few but Milley possessed. Strategic assurances to foreign military counterparts ensure that a constitutional transfer of power is peaceful by reducing the uncertainty such transfers of power create from the perspective of foreign partners. Milley's conduct hence serves as an example of how administrative capacity benefits from bureaucratic discretion to secure essential democratic goods.

## Responsible Public Service Norms

Individually, and collectively, these three cases illustrate the norms of responsible public service, that is, what makes it so that a bureaucrat's conduct constitutes them doing a *good job*. Across these three cases, different aspects stand out. For Petrov and the case of radon regulation, decisions based on expert expertise and access to broad evidence played a role. The case of General Milley was about maintaining constitutionality by reducing the risk of an escalatory cycle during the increased uncertainty of a democratic transfer of power. Both Petrov and Milley broadly contributed to maintaining political stability and public safety. In short, each of the three cases illustrated some aspects of the norms of responsible public service.

The norms of responsible public service are evident not only in the exemplary behavior of bureaucrats but also in the codes of ethics of public administration (ASPA 2020), and in the curricula of public administration programs, which are required to “lead and manage in the public interest” (NASPAA 2023). Moreover, the scholarship on public administration recognizes a long list of public values that inform public sector work (Bozeman 2007; Jørgensen and Bozeman 2007; Nabatchi 2018).<sup>6</sup> Taken together, such norms supply the standards by which bureaucratic discretion and administrative capacity are evaluated.

Although the specific content of these norms is contested, there is broad agreement in the literature on the roles of public administration, or the formal functions of having civil servants in a government. For example, authors agree that civil servants provide continuity during government transitions even if they may disagree on precisely what “provide continuity” involves. Setting aside the precise normative content of these formal functions, I submit that responsible public service is when bureaucrats play their role well. In other words, a

---

<sup>6</sup> These norms are incomplete or contested, however, in two respects. First, the precise content of these values and norms is unclear. Second, similarly the normative grounding of these norms—where they are coming from, or how they are justified—is contested. Yet, notwithstanding such controversies, there is significant agreement on the overall roles and functions of public administration in democracies.

bureaucrat does a good job when their conduct corresponds with the norms within the formal functions of public administration in a democracy. Starting with pre-theoretic examples of responsible public service may then be a guide to the content of these norms. The formal functions of public administration include the following (Heath 2020):

- (a) *Stability*. Public administration ensures political stability and a continuous performance of public service by compensating for the disfunction of the legislature. When a parliament fails to deliberate and enact legislation on pressing matters, public administration develops and issues regulations instead (Meier 1997).
- (b) *Long-term planning*. In transcending the limited time horizons of elections, public administration ensures fiscal and/or environmental sustainability.
- (c) *Consistency over time*. Public administration is often a moderating force against overcorrecting across different elected administrations, counteracting the “zigzag” of politics.
- (d) *Evidence*. Public administration promotes evidence-based decision making by
  - a. maintaining statistical agencies, assessing data needs, and collecting and processing administrative data;
  - b. developing and providing evidence-based policy advice.
- (e) *Public interest*. Public administration works in the public interest in various ways, for example, by:
  - a. promoting a technocratic standard of “good policy” (e.g., health, education, or general welfare) on issues that garner insufficient public interest;
  - b. advancing the interests of groups that fail to organize and articulate these interests themselves (because they are too dispersed or lack material resources);
  - c. limiting local, particularized, or concentrated highly organized particular interests.
- (f) *Resisting democratic pandering*. Public administration moderates the influence of whipped-up resentments against minorities, migrants, taxation, political opponents, among others.

(g) *Maintaining constitutionality.* Public administration checks both executive and legislative power. Public administration hence complements a similar function played by the judiciary.<sup>7</sup>

These functions, and executing them well, mean that a public administrator is doing a good job and not just performing a task well. Expertise, good evidence, maintaining stability and serving the public interest are values or norms that spell out what role public administrators should play in addition to what they are explicitly tasked with. The norms of responsible public service are the criteria by which to evaluate whether a public servant, even when they are not doing what they are formally supposed to do, are still doing a good job.

These roles, and hence the norms of responsible public service, enhance the quality of democracy. They ensure state capacity and government functioning, as discussed above, which increases institutional legitimacy. Moreover, responsible public service enhances the quality of democracy because responsible public service can promote equal representation. On the one hand, public administration can represent the voice of groups that are otherwise insufficiently heard. Public administration can facilitate deliberation. On the other hand, public administration can limit the influence of voices that are already heard particularly well. Public administration can be a corrective to the fact that those with money and time have greater means to coordinate and articulate their interests and influence decisions. In this role, public administration supports the fundamental democratic idea that every citizen should enjoy an equal opportunity to influence decisions of common concern. Several approaches to public administration stand in this tradition, specifically, the traditions known as New Public Administration and New Public Service (Waldo 2017; Denhardt and Denhardt 2000). Finally, public administration supports deliberation by providing evidence, for example, through official statistics, administrative data, and research funding. Insofar as the public interest is constituted by informed deliberations, providing evidence and expertise enhances the quality of democracy.

In some ways, these roles and the norms of responsible public service are controversial. The roles cited above see public administration as part of the political process. The picture of democracy that stands behind this idea of responsible public service norms is a picture of mixed constitutions, so that each branch of government contributes to democracy in a different manner. In this respect, the norms conflict with an alternative approach to public administration, known as the classical approach, that sees bureaucrats as cogs in a machine. In this alternative perspective, the role of public administration is simply to implement policy

---

<sup>7</sup> With one difference being that public administrators can do so pro-actively or in “real time.” By contrast, judicial action occurs after the fact.

and not to be involved in making it. Insofar as the classical approach to public administration is a reasonable view, these norms of responsible public service are controversial.

However, in important respects these functions should be relatively uncontroversial. Specifically, the above-cited norms of public administration aim to be politically neutral.<sup>8</sup> Citizens who see each other as free and equal should be able to accept these norms. The norms cannot further partisan aims nor clearly favor some others. The norms should be common constitutional ground. As such, the status of these norms is like that of the right to free expression or the right to vote. These are ideas that everyone who accepts the basic premises of democracy (mutual recognition as free and equal) should grant one other.

Alternatively, the norms can be seen as a requirement that allows democracy to work in practice. Bureaucrats come in whenever it is not feasible for every citizen or every legislator to be an expert, focus on the long term, and fully participate in the political process. Even the U.S. Congress, traditionally averse to such discretion, has regularly authorized federal agencies to spell out key provisions of legislation. Are SUVs passenger automobiles or light trucks? Are e-cigarettes drugs, devices, or tobacco products? Answers to these questions are political but such answers are given during policy implementation when the attention of the legislators has migrated elsewhere. In short, although the role of public administration is political, it's still democratic in an uncontroversial way: public administration should help every citizen to have an opportunity to influence decisions of common concern, whenever citizens can't do so themselves, and whenever the public interest can be promoted in some sense.

Yet, of course, what I just called “common constitutional ground” is contested—and increasingly so as competitive authoritarianism spreads around the globe. This much is clear from the fact that General Milley's case garnered so much controversy, including precipitating congressional hearings. The controversy surrounding General Milley seems to be driven by a populist perspective of democracy, one in which the authority of the legislative branch and elected officials should be absolute. These popular elements are not to be questioned by *unelected* public servants.<sup>9</sup> Thus, the common ground on which the norms of responsible civil service stand are not sociological facts but, rather, normative aspirations. The controversy surrounding General Milley's case can be seen as highlighting the extent to which some parts of the democratic system in the U.S. has deviated from this common ground.<sup>10</sup>

---

<sup>8</sup> Although what this means exactly can also be difficult to determine (Heath 2020, 33–40).

<sup>9</sup> Moreover, legislators could engage in strategic behavior to increase their own public profile as a defender of exactly this idea of a populist democracy.

<sup>10</sup> Insofar there ever was, in fact, such a common ground in practice (that there may be in theory).

## Practical Upshots

Bureaucrats sometimes can do a good job despite failing, to some extent, to perform their designated task. The three cases illustrated this difference between task and job. The roles and norms of responsible public service above mark the difference between performing a task well and doing a good job. We can now see more clearly how AI is a threat to democracy: Both matter for democracy, high task performance as well as doing a good job. Task performance is easy to measure, doing a good job is not. In building AI for government, the risk is that AI will be optimized for task performance, but not for doing a good job. If those who today play the role of Stanislav Petrov, Mark Milley, or the radon-regulation-advancing EPA administrators were replaced by AI agents, we would see a loss in democratic quality. AI would perform the tasks, but unlike the example above, would not know what it means to do a good job. When we evaluate administrative capacity but evaluate administrative capacity only narrowly as the performance of tasks, we fail to capture how bureaucratic discretion can support democracy.

### AI for Tasks but not Jobs

AI threatens democracy insofar as it is easier for an AI to perform a task well than it is for it to do a good job. That AI may not be as good as human bureaucrats at doing a good job is not primarily an engineering problem. Even when we figure out the technology, it's still hard for AI to uphold the norms of responsible public service.

For one, the norms of responsible public service are far from self-evident. As such, we have an epistemological problem: It's not clear what makes for good public service. Moreover, responsible public service is a normative profession. Doing a good job in public administration requires one to answer dilemmas and resolve values that conflict practically with one another. For example: Is the additional burden of crafting more complex building codes worth the increase in public health that we expect from regulating radon gas? Is General Milley serving primarily the President or the American people? AI may be able to engage in moral reasoning, but it's not clear what exactly it takes for civil servants to make such tradeoffs well.

Most immediately, however, the biggest risk is that the difference between doing a task well and doing a good job is not as widely recognized as it should be. Bureaucrats work in the background. Their role is difficult to explain to an audience that thinks that "democracy" essentially means "majority rule" or "rule by the people." As a result, when bureaucrats are replaced or augmented by AI, then, when we fail to see what they contribute to democracy, we might build a replacement for the *task* that the bureaucrats did, but *not* replace what would make them good at their job. We end up replacing bureaucrats with AI, but only the part that's easy to replace.

## Evidence and Validation Standards

The norms of responsible public service offer practical guidance and a positive vision for governing AI in the public sector. These norms both justify some existing policies and suggest new approaches to protect democratic functioning as administrative decisions become increasingly automated. The norm of maintaining good evidence becomes especially critical when evidence is not only used for policy analysis but also as training data for AI systems that drive administrative decisions.<sup>11</sup>

The 2021 U.S. Executive Order on *Advancing Racial Equity* exemplifies this norm of good evidence. This order requires that federal datasets be disaggregated “by race, ethnicity, gender, disability, income, veteran status, or other key demographic variables” since the existing “lack of data ... impedes efforts to measure and advance equity” (Executive Office of the President 2021). This is because aggregate data can obscure inequities and lead to discriminatory outcomes when used in automated systems.

One aim of this Executive Order, expanding the range of variables that are collected in administrative data, should hence be uncontroversial. Gathering good data is a core formal function of public administration (item d above). Unfortunately, the label “racial equity” makes this order more controversial than its specific content. This specific policy content—to disaggregate administrative data—can be derived from the much less controversial idea that good evidence is necessary for good public administration. The title on “racial equity,” in contrast, suggests preferential treatment to mitigate historic injustices—a far more controversial idea. Accordingly, the new Trump administration quickly revoked this order.

The norm of responsible public service can also support AI model validation requirements. Models that classify pictures of faces by demographic characteristics should be benchmarked in a way that disaggregates accuracy by race and gender (Buolamwini and Gebru 2018). This is a demand of responsible public service that can, in turn, be grounded in broader democratic ideas of equality or transparency.

One case involving AI where the standards of evidence seem to have been violated is the MiDAS scandal in Michigan. Michigan developed an automated fraud detection system, MiDAS. From 2013 to 2015, the system issued thousands of erroneous penalties, falsely accusing at least 20,000 people of fraud, at a rate of reportedly over 90 per cent of those accused (Felton 2016; Charette 2018). The high error rate stemmed from data import problems: information was corrupted or lost during migration from the legacy system. MiDAS made fraud determinations without access to complete data (Egan 2017). These problems were only discovered after MiDAS was deployed without adequate validation. This

---

<sup>11</sup> At the same time, what counts as “good” evidence is not always clear.

violated both the norm of good evidence and the norm of serving the public interest. Those wrongly accused of fraud lacked political representation to contest the system's deployment, thus making it the responsibility of public administrators to ensure proper safeguards. However, there was no avenue for public administrators to play this role.

This suggests a key principle for AI governance: evidential standards should scale with the stakes of deployment. High-stakes applications, such as financial penalties for suspected benefit fraud, require rigorous validation before deployment. When such validation proves infeasible, the stakes should be lowered. For instance, models could be deployed in advisory rather than determinative roles. Alternatively, models could be deployed in a determinative role but with an initial trial phase during which penalties are reduced. In the case of MiDAS, meaningful validation would have required human review of a representative sample of fraud determinations, verification that data migration preserved information integrity, and testing on cases with known outcomes. This approach would protect vulnerable populations that often lack representation while preserving public trust in both AI systems and government institutions.

### Preserving Administrative Discretion Through Model Governance

The cases of Petrov and Milley highlight how public servants use their positions to maintain stability and resist pressures that threaten the public interest. This capacity depends on employment protections, professional expertise, and zones of discretion that allow civil servants to exercise judgment. Yet, AI adoption threatens these mechanisms through the centralization and automation of decision-making (Zouridis et al. 2020).

AI continues a development that started many decades ago when the introduction of “information communication technologies” replaced street-level with screen-level discretion — and the administrative power of many caseworkers was centralized in organizational processes (Bullock 2019; Bovens and Zouridis 2002). The adoption of AI in the administrative state hence advances the centralization of administrative power that, at the same time, closes the usual channels through which the norms of responsible public service were previously upheld.

One important factor, among others, that enables mechanisms of administrative discretion is that civil service jobs tend to enjoy significant protections. One response to the erosion of human administrative discretion could be to replicate some of the mechanisms that allowed human bureaucrats to exercise administrative discretion to models.

Specifically, we could establish model deployment rules that function analogously to civil service protections. Just as civil servants cannot be arbitrarily dismissed when administrations change, AI models in production should not be subject to immediate wholesale revision by

newly elected officials. Such restrictions would not only prevent the rapid politicization of automated decisions but also ensure that model changes undergo proper testing and validation. These deployment rules would help maintain the stability and consistency that bureaucrats traditionally provide across electoral cycles.

We could establish model deployment rules that replicate civil service job protections for AI models. Specifically, there should be a minimum period before models that are deployed in production cannot be revised or replaced. Such limits could not only prevent swift and blanket changes of models and decision-procedures by newly elected officials, but they might ensure that these revisions are sufficiently tested. A newly elected officeholder should not be able to overhaul immediately the entire universe of models in his or her purview.<sup>12</sup>

Unfortunately, blanket rules against model replacement won't do. Sometimes models need to be changed or replaced rapidly. A pandemic, an economic crisis, or a major policy change, so-called "distribution shifts," alter the data environment, making models trained on historical patterns unreliable. The governance challenge lies in distinguishing such legitimate technical grounds for replacing models from political interference.

### Maintaining Channels for Professional Judgment, Dissent, and Innovation

As AI systems create comprehensive audit trails and standardize decision processes, they risk eliminating the informal channels through which civil servants have historically exercised professional judgment and raised concerns. Petrov's ability to override Oko depended not just on his expertise but on organizational structures that permitted such discretion. Similarly, bureaucratic initiatives like the EPA's radon regulation emerged from spaces where civil servants could identify and pursue emerging issues independently, often pushing the limits of their authorized mandate.

To preserve these functions of expertise and advancing the public interest, AI governance should explicitly maintain channels for bureaucratic discretion, dissent, and innovation. This might include formal processes for civil servants to flag concerns about automated systems, and protected spaces for bureaucratic innovation outside algorithmic optimization. Whistleblower protections should extend to those who identify problems with AI systems, and organizations should establish clear escalation paths for cases where human judgment diverges from algorithmic recommendations.

---

<sup>12</sup> Blanket "job protection" for models would not do, however. Sometimes models need to be updated quickly (e.g., because of what is known as distribution shift), and model deployment rules that put minimum lifetimes on models would make such updates impossible.

## Limits to Automation

Finally, the norms of responsible public service suggest that some governmental functions should resist automation entirely. If the use of AI significantly weakens mechanisms that enable public servants to do a “good job,” then AI should probably not be adopted. Where AI cannot replicate the holistic reasoning, moral judgment, or democratic representation that bureaucrats routinely provide, automation should be approached with extreme caution or avoided altogether. Judges who consider the full context of individual cases, social workers who advocate for vulnerable populations, and intelligence analysts who integrate disparate forms of evidence all perform roles where the difference between task and job remains particularly pronounced.

## Concluding Remarks

The challenge will be to devise clever ways of augmenting public administration with AI. This is a challenge for researchers, administrative law, and government reformers. Humans and AI have complementary strengths. Some of these strengths are clear. For example, judges have access to a broader range of evidence than AI. In short, humans are good in processing data that is *wide*, whereas AI is good at processing data that is *deep* (Ludwig and Mullainathan 2021). Similarly, the human contribution to such AI-augmented administrative decision making could be to uphold the norms of responsible public administration.

However, some of the strengths that human decision makers can bring to a human–AI collaboration are not sufficiently recognized by government reformers. Specifically, how bureaucrats contribute to democratic legitimacy is often overshadowed by a populist conception of democracy as simple majority rule. Yet, public servants can exercise bureaucratic discretion to support democratic legitimacy. This is often overlooked when we evaluate state capacity: Cultivating independent expertise, responding quickly to emerging public health challenges, and checking executive power—these are elements of responsible public service that AI may jeopardize.

Through the cases of Petrov, the EPA’s radon regulation, and General Milley this chapter illustrated that: public servants do more than execute tasks—they uphold norms of responsible public service that contribute to democracy; and bureaucrats safeguard stability, gather and maintain evidence, pursue the public interest, resist democratic pandering, and maintain constitutional boundaries. These functions represent not a betrayal of democratic principles, but instead their fulfillment through complementary institutional mechanisms. Public servants exercise bureaucratic discretion as a democratic practice.

AI poses a distinct threat to these democratic practices because it excels at performing tasks while struggling to replicate what makes a job well done. The risk is not merely technical,

that is, AI cannot yet reason morally or integrate evidence holistically—but conceptual and normative. Normatively, whether bureaucrats should have the kind of power that enables them to uphold the norms of responsible public service is contested. Conceptually, when we fail to recognize the difference between performing a task well and doing a good job, we build systems that optimize the wrong objectives. We displace what bureaucrats do, but not what makes their work valuable to democracy.

Norms of responsible public service provide more than cautionary insights. They can also offer constructive guidance for AI governance. These norms are more comprehensive than narrow concerns about algorithmic fairness, yet more grounded in a constitutional consensus than contested theories of substantive justice. They reveal why existing policies like data disaggregation requirements matter, suggest new approaches like model deployment protections, and identify where automation should be limited or avoided.

The path forward requires making the implicit explicit. The professional norms that guide responsible public service, long taken for granted, must be articulated clearly enough to shape AI development and deployment. Standards by which state capacity and the exercise of bureaucratic discretion are evaluated need to incorporate the idea of what it means for public servants—human or AI—to do a good job. I argued in this chapter that there is a distinction between performing a task and doing a good job and that norms of responsible public service make the difference. But the content of these norms and how they are normatively grounded is a research project for another day.

The stakes extend beyond the public sector. The erosion of professional norms in favor of task optimization threatens expertise and judgment across policy making domains. But it is in government where the stakes are highest, where the difference between task and job most directly impacts democratic values, and where the lessons for the responsible deployment of AI are most urgently heeded. By recognizing how public servants who do a good job contribute to state capacity and democracy, we can better navigate the transformation that AI brings to the administrative state and democracy generally.

## References

- Allen, Danielle, and E. Glen Weyl. 2024. "The Real Dangers of Generative AI." *Journal of Democracy* 35 (1): 147–62.
- ASPA. 2020. "Code of Ethics." American Society for Public Administration: Code of Ethics. <https://www.aspanet.org/ASPA/Code-of-Ethics/Code-of-Ethics.aspx>.
- Bengio, Yoshua. 2023. "AI and Catastrophic Risk." *Journal of Democracy* 34 (4): 111–21. <https://doi.org/10.1353/jod.2023.a907692>.
- Bovens, Mark, and Stavros Zouridis. 2002. "From Street-Level to System-Level Bureaucracies: How Information and Communication Technology Is Transforming Administrative Discretion and Constitutional Control." *Public Administration Review* 62 (2): 174–84. <https://doi.org/10.1111/0033-3352.00168>.
- Bozeman, Barry. 2007. *Public Values and Public Interest: Counterbalancing Economic Individualism*. Georgetown University Press.
- Bullock, Justin B. 2019. "Artificial Intelligence, Discretion, and Bureaucracy." *The American Review of Public Administration* 49 (7): 751–61. <https://doi.org/10.1177/0275074019856123>.
- Buolamwini, Joy, and Timnit Gebru. 2018. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, January 21, 77–91. <https://proceedings.mlr.press/v81/buolamwini18a.html>.
- Charette, Robert N. 2018. "Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold - IEEE Spectrum." *IEEE Spectrum*, January 24. <https://spectrum.ieee.org/riskfactor/computing/software/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold>.
- Denhardt, Robert B., and Janet Vinzant Denhardt. 2000. "The New Public Service: Serving Rather than Steering." *Public Administration Review* 60 (6): 549–59. <https://doi.org/10.1111/0033-3352.00117>.

- DHS. 2025. “United States Citizenship and Immigration Services – AI Use Cases | Homeland Security.” February 24. <https://www.dhs.gov/ai/use-case-inventory/uscis>.
- Egan, Paul. 2017. “Data Glitch Was Apparent Factor in False Fraud Charges against Jobless Claimants.” *Detroit Free Press* (Lansing), July 30. <https://www.freep.com/story/news/local/michigan/2017/07/30/fraud-charges-unemployment-jobless-claimants/516332001/>.
- Eggers, William D., and Thomas Beyer. 2019. “AI-Augmented Government.” In *Government Trends 2020*. Deloitte Insights. Deloitte Development LLC. <https://www2.deloitte.com/us/en/insights/industry/public-sector/government-trends/2020/ai-augmented-government.html>.
- Engstrom, David Freeman, Daniel E. Ho, Catherine M. Sharkey, and Mariano-Florentino Cuéllar. 2020. “Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies.” *NYU School of Law, Public Law Research Paper No. 20-54*, ahead of print, February 1. <https://doi.org/10.2139/ssrn.3551505>.
- Executive Office of the President. 2021. “Advancing Racial Equity and Support for Underserved Communities Through the Federal Government.” Federal Register, January 20. Federal Register 7009. <https://www.federalregister.gov/documents/2021/01/25/2021-01753/advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government>.
- Federal Chief Information Officers Council. 2023. “Federal AI Use Cases 2023 (AI.Gov).” CSV. [https://ai.gov/wp-content/uploads/2023/10/2023%20Consolidated%20AI%20Use%20Case%20Inventory%20\(PUBLIC\).csv](https://ai.gov/wp-content/uploads/2023/10/2023%20Consolidated%20AI%20Use%20Case%20Inventory%20(PUBLIC).csv).
- Felton, Ryan. 2016. “Michigan Unemployment Agency Made 20,000 False Fraud Accusations – Report.” US News. *The Guardian*, December 18. <https://www.theguardian.com/us-news/2016/dec/18/michigan-unemployment-agency-fraud-accusations>.
- Glaze, Kurt, Daniel E. Ho, Gerald K. Ray, and Christine Tsang. 2024. “Artificial Intelligence for Adjudication: The Social Security Administration and AI Governance.” In *The Oxford Handbook of AI Governance*, edited by Justin Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M Hudson, Anton Korinek, and

- Baobao Zhang. Oxford University Press.  
<http://academic.oup.com/edited-volume/41989/chapter/355439450>.
- Goldberg, Jeffrey. 2023. "The Patriot." Politics. *The Atlantic*, September 21.  
<https://www.theatlantic.com/magazine/archive/2023/11/general-mark-milley-trump-coup/675375/>.
- Hankins, Emma. 2023. "Release: 2023 Government AI Readiness Index Reveals Which Governments Are Most Prepared to Use AI." Oxford Insights, December 6.  
<https://oxfordinsights.com/insights/release-2023-government-ai-readiness-index-reveals-which-governments-are-most-prepared-to-use-ai/>.
- Harrison, Kathryn, and George Hoberg. 1991. "Setting the Environmental Agenda in Canada and the United States: The Cases of Dioxin and Radon." *Canadian Journal of Political Science/Revue Canadienne de Science Politique* 24 (1): 3–27.  
<https://doi.org/10.1017/S0008423900013391>.
- Heath, Joseph. 2020. *The Machinery of Government: Public Administration and the Liberal State*. Oxford University Press.
- Himmelreich, Johannes. 2023. "Should We Automate Democracy?" In *The Oxford Handbook of Digital Ethics*, edited by Carissa Véliz. Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780198857815.013.33>.
- Jørgensen, Torben Beck, and Barry Bozeman. 2007. "Public Values: An Inventory." *Administration & Society* 39 (3): 354–81.  
<https://doi.org/10.1177/0095399707300703>.
- Kreps, Sarah, and Doug Kriner. 2023. "How AI Threatens Democracy." *Journal of Democracy* 34 (4): 122–31. <https://doi.org/10.1353/jod.2023.a907693>.
- Lewis, Michael, Casey N. Cep, Dave Eggers, et al., eds. 2025. *Who Is Government? The Untold Story of Public Service*. Riverhead Books.
- Lubold, Gordon. 2021. "Mark Milley Says Calls to Chinese General Were Within His Duties." Politics. *Wall Street Journal*, September 17.  
<https://www.wsj.com/politics/national-security/mark-milley-says-calls-to-chinese-general-were-within-his-duties-11631878596>.

- Ludwig, Jens, and Sendhil Mullainathan. 2021. "Fragile Algorithms and Fallible Decision-Makers: Lessons from the Justice System." *Journal of Economic Perspectives* 35 (4): 71–96. <https://doi.org/10.1257/jep.35.4.71>.
- Meier, Kenneth J. 1997. "Bureaucracy and Democracy: The Case for More Bureaucracy and Less Democracy." *Public Administration Review* 57 (3): 193. <https://doi.org/10.2307/976648>.
- Nabatchi, Tina. 2018. "Public Values Frames in Administration and Governance." *Perspectives on Public Management and Governance* 1 (1): 59–72. <https://doi.org/10.1093/ppmgov/gvx009>.
- NASPAA. 2023. "NASPAA Accreditation Standards." Commission on Peer Review and Accreditation, October 13. <https://www.naspaa.org/sites/default/files/docs/2024-01/NASPAA%20Accreditation%20Standards%20-%202024%20FINAL%20with%20rationale.pdf>.
- O’Leary, Rosemary. 2019. *The Ethics of Dissent: Managing Guerrilla Government*. CQ Press.
- Waldo, Dwight. 1955. *The Study of Public Administration*. Doubleday.
- Waldo, Dwight. 2017. *The Administrative State: A Study of the Political Theory of American Public Administration*. Routledge.
- Zouridis, Stavros, Marlies van Eck, and Mark Bovens. 2020. "Automated Discretion." In *Discretion and the Quest for Controlled Freedom*, edited by Tony Evans and Peter Hupe. Springer International Publishing. [https://doi.org/10.1007/978-3-030-19566-3\\_20](https://doi.org/10.1007/978-3-030-19566-3_20).