

NISKANEN C E N T E R

THE *HOW* WE NEED NOW: A CAPACITY AGENDA FOR 2025 AND BEYOND

Jennifer Pahlka

Andrew Greenway

December 2024

The Niskanen Center is a 501(c)3 issue advocacy organization that works to change public policy through direct engagement in the policymaking process.

NISKANEN CENTER | 1201 New York Ave NW | WASHINGTON, D.C. 20005
www.niskanencenter.org | For inquiries, please contact ltavlas@niskanencenter.org

Contents

Part 1: Introduction	2
Part 2: Fixing government, then and now	4
Part 3: Losing our way	6
The brake is stuck	7
The steering is loose	9
We're trying to go everywhere at once	12
We've abdicated responsibility for navigating	14
Part 4: Finding our way	18
1. Hire the right people and fire the wrong ones	18
Reform the civil service	18
Rebalance the brake and the gas	20
2. Reduce procedural bloat	21
Take on policy, procedure, and regulatory accretion	22
Principles for de-proceduralization	22
Right-size procedures that impact every agency	23
Reduce policy accretion	23
Reduce the surface area for legal attack	24
Resist trade-off denial	25
3. Invest in digital infrastructure	26
Transition to product model funding	26
Invest in centralized and decentralized teams	30
Invest in platforms	31
Right-size AI guardrails	32
4. Close the loop around outcomes	33
Align funding to support feedback loops	36
Fund internal competence first	38
Align oversight to new models	38
Focus on enablement over mandates and constraints	40
Invest in Congressional capacity	42
Incentivize legislators to close the loop	43
Make new operating models the norm	43
Conclusion: What a different <i>how</i> offers	44
Part 5: Summary of recommendations	45

Consider these recent headlines:

- **“The U.S. Navy Can’t Build Ships,”** *Foreign Policy*, May 2024.
- **“New report estimates U.S. fraud losses exceed \$233 billion annually,”** *The Washington Post*, April 2024.
- **“Tens of thousands of Afghans applying for US visas still face major delays,”** *CNN*, July 2022.
- **“The U.S. ‘Fast-Tracked’ a Power Project. After 17 Years, It Just Got Approved,”** *Wall Street Journal*, May 2023.
- **“How the VA Fails Veterans on Mental Health,”** *ProPublica*, January 2024.¹

These outcomes are not the result of lack of diligence by our nation’s public servants. Our administrative state’s structures, processes, and ways of working are simply no longer fit for the jobs we need them to do. Both our airwaves and discourse on Capitol Hill are overwhelmed with debates about policy. What our government should do sets the terms of debate. But the *what* matters little if our nation’s leadership doesn’t put equal consideration into *how* to do it.

Government can and must be reorganized to meet the current moment. This requires prioritizing the government’s ability to achieve its policy goals – what academics call “state capacity” – in a way not seen since the early 1900s. It’s time.

This paper provides a thorough diagnosis of today’s state capacity challenges, as well as a promising path forward. Transforming our government for the modern era is a long-term effort, but we must start now.

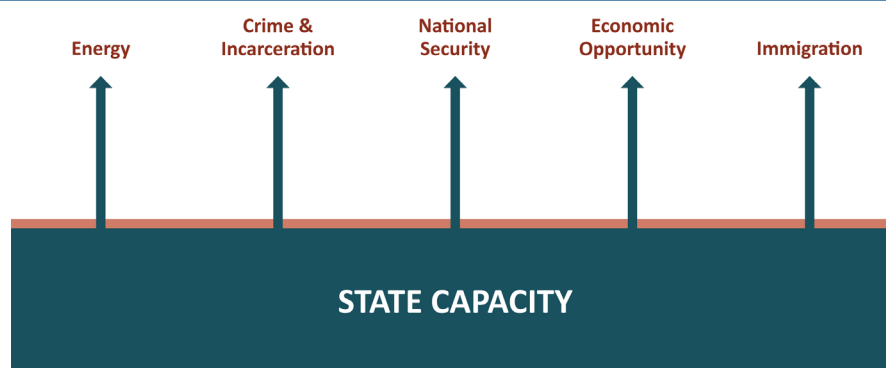
1. Gil Barndollar and Matthew C. Mai, “The U.S. Navy Can’t Build Ships,” *Foreign Policy*, May 17, 2024, <https://foreignpolicy.com/2024/05/17/us-navy-ships-shipbuilding-fleet-china-naval-race-pacific/>. Tony Romm, “New report estimates U.S. fraud losses exceed \$233 billion annually,” *The Washington Post*, April 16, 2024, <https://www.washingtonpost.com/business/2024/04/16/government-fraud-gao/>. Jennifer Hansler et al., “Tens of thousands of Afghans applying for US visas still face delays,” *CNN*, July 18, 2022, <https://www.cnn.com/2022/07/18/politics/afghans-us-visas-delays/index.html>. Jennifer Hiller and Andrew Restuccia, “The U.S. ‘Fast-Tracked’ a Power Project. After 17 Years, It Just Got Approved,” *The Wall Street Journal*, May 18, 2023, <https://www.wsj.com/articles/the-u-s-fast-tracked-a-power-project-after-17-years-its-nearing-approval-1a7edb86>. Kathleen McGrory and Neil Bedi, “How the VA Fails Veterans on Mental Health,” *ProPublica*, January 9, 2024, <https://www.propublica.org/article/how-veterans-affairs-fails-mental-health-patients>.

Part 1: Introduction

President Trump and the 119th Congress will inherit profound challenges: an unstable global order, massive disruptions from advanced technology, crumbling infrastructure, rising healthcare costs, and an unprecedented budget deficit, just to name a few. At the same time, the U.S. government stands at the helm of the largest public budget on the planet and a set of institutions created to address these challenges in the service of the public good. In steering these incredible resources, it is charged with answering difficult and enduring questions: What strategies will keep Americans safe and secure, at home and overseas? What investments are necessary to make this an increasingly prosperous nation, and who should be making them? What must be done to make sure citizens have access to quality healthcare, education, employment, and housing? Who should pay how much in tax to fund all this?

Yet a question all too rarely asked is: How well do our institutions actually work in addressing these and other questions? Can they translate the political intent – of any party – into public impact? Once we decide *what* our policies should be, do we know *how* to implement them in real life?

Every policy goal relies on the capacity of the state to implement



For Republican and Democratic leaders alike, the impact of their time in office will rest upon the government's capacity to successfully implement the policies they devise and enact. The prospect of piloting a government critically low on state capacity poses a deep problem for either party. That is true not only when it comes to establishing a policy legacy. Our politics have become so chaotic in part because the public has lost trust in the government's ability to deliver on its promises. Restoring this public trust is both a moral and a political imperative for leaders who want to leave their mark.

The time is now also right for a focus on effective government because any fiscal margin for error we might have enjoyed earlier in this century has evaporated. In the last eight years, the Trump and Biden administrations have committed to unprecedented public investments for economic recovery, infrastructure, and national competitiveness. Meanwhile, borrowing costs have risen steeply and obligations to our aging population like Social Security and Medicare have grown, even as domestic and global challenges will create new fiscal needs.

We believe in the old adage about making the most of a crisis. In that respect, the pandemic was something of a missed opportunity. It forced our government to work differently and meet needs faster. But as the sting of crisis abated, things slid back towards the status quo. Now, the constraints facing the next administration provide an unmissable opportunity to change how government works for the better. But we need to be committed to turning this urgency into lasting change, and applying it to slow-boiling problems, too.

Rebuilding our state capacity — the *how* of our government — is an arduous task that requires a stomach for learning and fixing the rules and culture of the bureaucracy. But make no mistake: whether or not we do it is not a technocratic choice. It's a political choice — a leadership choice. Previous administrations may not have actively set out to make dealing with public services a burden, or to cripple our nation's military readiness with policy clutter amounting to hundreds of copies of War and Peace, or create a veto-crazy that limits our ability to build infrastructure. But neither have they made fixing it a priority. This one can, and should.

Underneath the sound and fury of politics, hundreds of thousands of civil servants who form the backbone of our government, serving through administrations of both parties, are going about their job of keeping the country going. Many are comfortable with the status quo, for sure, but a growing band is hungry for permission to reinvent the bureaucracy in pursuit of the mission that inspired them to serve in the first place. These public sector patriots want to make government work better, to have a firm and capable grip on the problems it is charged with solving. They are an enormous national resource; the incoming leadership would do well to recognize them and leverage their energy to the fullest extent.

Some agents of change are successful, thanks to unusual circumstances that created political cover for them to do things differently. They have been given the space to apply new ways of working, often driven by an emergency — and they delivered. Yet these exceptions have remained just that — exceptional. When the change agents move on or the exigency of the crisis they are finally empowered to solve subsides, the magic often disappears. With greater political will across the executive and legislative branches, we can make a culture of experimentation and improvement the norm rather than a welcome surprise.

Our two parties (and even the factions within them) often speak about state capacity in very different languages and have differing priorities for where that muscle should be deployed. But the vast majority of elected leaders on both sides of the aisle would like the bureaucracy to be capable of achieving the missions it is assigned. This report is a guide to regaining that capability. Though we focus on the federal government here, the report draws on lessons from all levels of U.S. government and from countries around the world that are struggling with similar challenges. We take a particular look at the UK, home of one of the authors, for inspiration about how to move forward.

We first look at both structural and cultural drivers of incapacity and then turn our attention to prescriptions. We offer four priority areas for building state capacity — the *how* of our government — over the coming decade, and detail what reformers must do to enable each of them. While the picture we paint of our government's capacity today is not always pretty, we believe that achieving a more responsive, effective, and confidence-inspiring government is a matter of will. Putting capacity at the forefront will help even a highly-constrained administration to reduce risk, save money, and deliver what the American people need. It will enable government to address the complex problems we have failed to unravel for decades and to face the new challenges this disruptive century continues to bring.

As a country, we will continue to disagree about what the government should do. But we can, and must, agree that it must be more capable of doing the core things it sets out to do. Getting there will take more than new policies. Now is the time to fix our *how*.

“I’m in year 34 of Federal civilian service. I’ve never before seen the kind of bureaucratic squeamishness I’ve seen in the last 5 years. There’s so much fear of making a mistake and getting embroiled in a lawsuit that people are maximally risk-averse. This is murdering the mission.”

- Online post from civil servant

Part 2: Fixing government, then and now

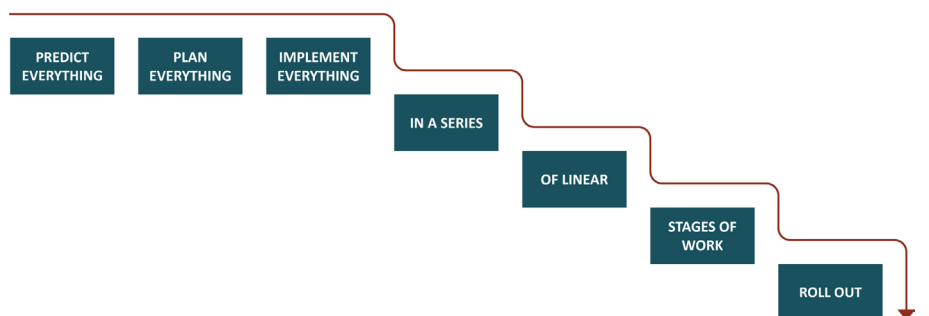
For some, the task of fixing the government seems almost impossible. To illustrate how we might succeed in reforming government for the modern day, it is instructive for us to briefly revisit America's last major step forward in government reform: the Progressive movement of the 1890s to the 1920s.

Progressives bridged the Republican and Democratic parties, and they focused on the *how*. Prior to the reforms they championed, government was an ad hoc, unprofessionalized set of institutions that ran on what we today would call graft. The Progressives saw themselves as heirs to the reformers who had finally managed, in 1883, to establish a merit-based system for federal employment and create a Civil Service Commission to oversee it. They sought to professionalize public administration by promoting training and standards for public servants, ensuring that government employees were competent and qualified. Progressives also advocated for standardization of procedures and greater transparency in government operations to reduce waste and corruption. This included standardized accounting practices, open meetings, and public records laws.

This history matters because it can serve as inspiration, example, and caution. Progressivism showed it was possible to build a bipartisan reform movement around a broad set of concerns that echoes those of today: a sense that economic developments had disrupted the social fabric, and that government was no longer fit for purpose. That consensus drove decades of major institutional reforms, especially at the state and local level, even though Progressives disagreed on critical issues, including corporate regulation and cultural matters such as alcohol regulation and immigration.

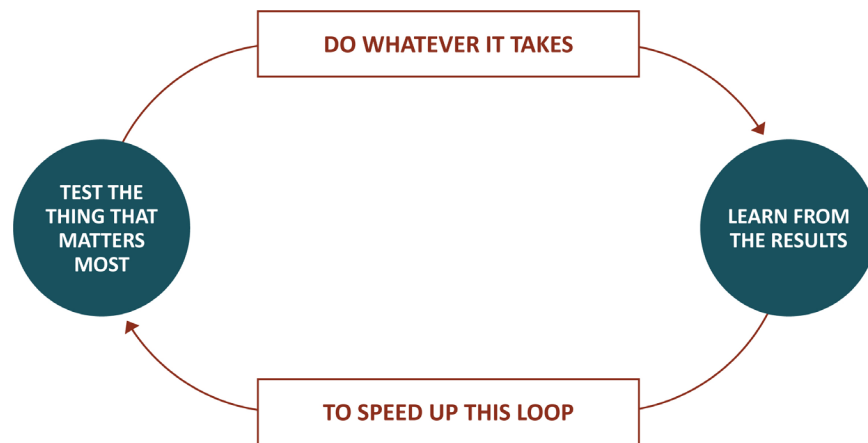
Institutionally, the Progressives logged many successes, but they are also in part responsible for making American government more complicated: They had a penchant for layering new structures on top of the old in order to work around corruption and inefficiency. We suffer mightily today from the accumulation of those layers. And while important pro-integration reformers counted themselves as Progressives, the movement was also deeply intertwined with racism and segregation.

This legacy does not provide us with straightforward lessons for the present, but it allows us to proceed with a sense for how today's challenges grow out of successes and failures of the past. There are cycles of government renewal, and we are urgently overdue for another big push. The Progressives fought to rebuild government fit for the Industrial Era. The government we have today operates largely along Industrial Age patterns, in ways that look like this:



2. Andrew Greenway and Tom Loosemore, "The Radical How," Public Digital, February 2024, <https://options2040.co.uk/wp-content/uploads/2024/02/The-Radical-How.pdf>.

This process, sometimes described as ‘waterfall’ in homage to the sequential, one-way steps it is supposed to follow, is typical of major government programs. State capacity advocates, drawing on liberal and conservative ideas, and working within both parties, are fighting for a government fit for the digital era - one that works more like this:



3

Government is taking steps to catch up with the times, in pockets. Digital services in the United States are getting better for constituents as diverse as veterans,⁴ international travelers,⁵ and low-income tax filers.⁶ Our intelligence infrastructure has provided critical insight into actions by adversaries. Weather predictions have gotten so accurate that almost every detail of Hurricane Helene’s track and intensity was predicted from early in the storm’s formation.⁷

Despite these bright spots, we have so much further to go. Travelers can now renew their passports online, but many students couldn’t file for financial aid last year and enrollment of 18 year-olds in the incoming freshman class declined 5% as a result.⁸ Our intelligence wins are overshadowed by grave concerns about our military readiness. Big data helps us foresee the weather, but we failed to convince many residents to evacuate away from the hurricane, and left others without power and water for weeks. There is success to build on, but much building ahead.

And, of course, even where we are catching up, the pace of change is accelerating. So much has been said elsewhere about the speed of AI advancements that we won’t belabor the point. The communications revolution also means social and global conditions are changing faster every day. Yet we’re barely moving the majority of government from the Industrial Age to the digital one. We’re mostly still a full chapter behind in the book.

3. *Ibid.*

4. Amy Kluber, “VA Celebrates Digital Services Innovation to Boost Veterans Benefits Access,” GovCIO, July 2, 2023, <https://govciomedia.com/va-celebrates-digital-services-innovation-to-boost-veterans-benefits-access/>.

5. April Rubin, “Passport renewals go digital to ease wait times for American Travelers,” Axios, September 18, 2024, <https://www.axios.com/2024/09/18/online-passport-renewal-us-travelers>.

6. The Associated Press, “‘So darn easy’: IRS’ free Direct File pilot program already a hit with early users,” NBC News, March 6, 2024, <https://www.nbcnews.com/business/taxes/-darn-easy-irs-free-direct-file-pilot-program-already-hit-early-users-rcna142132>.

7. National Hurricane Center And Central Pacific Hurricane Center, Hurricane HELENE Advisory Archive <https://www.nhc.noaa.gov/archive/2024/HELENE.shtml>.

8. Jessica Dickler, “College enrollment falls 5% for 18-year-old freshmen; FAFSA failures to blame, experts say,” MSN, December 4, 2024, <https://www.msn.com/en-us/education-and-learning/higher-education/college-enrollment-falls-5-for-18-year-old-freshmen-fafsa-failures-to-blame-experts-say/ar-AA1vgpM0?ocid=BingNewsSerp>.

Every government around the world is in the same boat, but some are starting to plot different courses. The few whose institutions seem fit for the digital era (like Estonia and Singapore) are younger and smaller countries, less burdened by legacy and more adaptable. Of the rest, each has distinct challenges. But most are recognizing the increasingly urgent need to shed the yoke of the past and adapt.

We can take some inspiration from our friends across the pond, where both leadership and line staffers in the UK government are starting to question the suitability of the structures, practices, and culture of their public sector to the formidable tasks they face, and conclude that change is sorely needed. One of us co-authored a report for the new leadership about how to preserve the progress their government had already made in this transition and accelerate the momentum towards new and better ways of working.⁹ Interest in finding new approaches to the *how* of government is high.

To be sure, the UK government continues to face many of the same profound challenges as ours, and is very early on its journey to the capacity it needs, but it is at least a little bit ahead. The new government announced a £100m (\$125m) innovation fund in its first budget to “support the development of a new approach to improving public services,”¹⁰ investing in experimentation and test-and-learn teams. It also promised to “overhaul how recruitment is carried out across the Civil Service,”¹¹ as a starting point for wider thinking on other systemic changes. Our hope with this report is to spur the U.S. to catch up, not by copying the UK’s strategies and tactics, but by taking this agenda at least as seriously as it’s being taken in the UK.

In the 1860s and 1870s, the UK had its own reckoning with Progressivism, radically reforming a governmental system riven by patronage and graft around a newly merit-based bureaucracy. These reforms were inspired in part by the practices of the colonial Indian civil service, an entity which — like Estonia and Singapore of today — presented a relatively free field for innovation unburdened by legacy. These superior practices embarrassed the powers that were in London into reforms. And in turn, those reforms in the UK inspired Progressives in the United States to walk their own path towards rethinking how their government should do things. Now would be a good time for this historical cycle to return.

Part 3: Losing our way

American bureaucracy is vast and heterogenous, and solving a performance problem within any agency requires a deep dive into the particulars of the situation. Bureaucratic failure arises from a complex interaction of rules and culture, and anyone hoping to fix it must appreciate both. But certain patterns of dysfunction arise time and again. In this section, we highlight those patterns to give policymakers an aerial view of the landscape and set the table for our recommendations. In the next section, we hope to give them a map that leads through it, and indicates the most promising areas in which to plant a flag.

Today, America has a government that’s stuck. We can say where we need to go — onshoring manufacturing, defending against new threats, deepening our economic advantages — but too often, we can’t seem to get

9. Greenway and Loosemore, “Radical How,” <https://options2040.co.uk/wp-content/uploads/2024/02/The-Radical-How.pdf>.

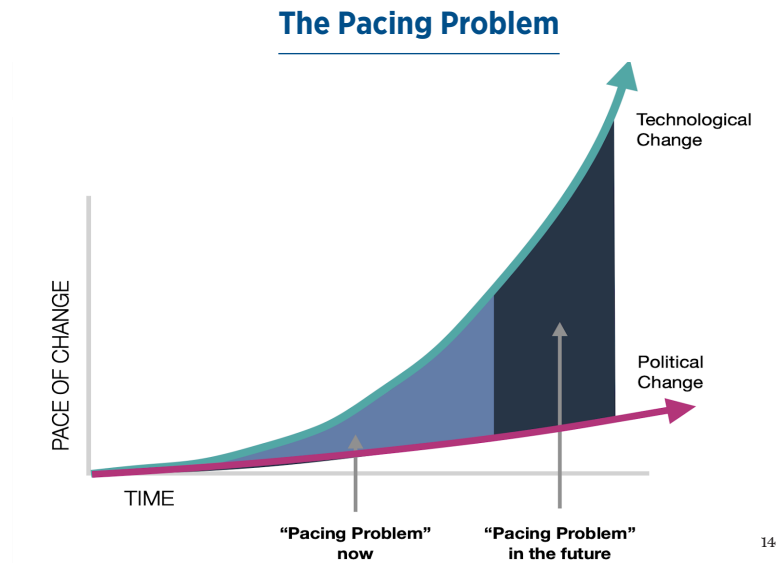
10. “Autumn Budget 2024: Fixing the Foundations To Deliver Change,” HM Treasury, October 30, 2024, https://assets.publishing.service.gov.uk/media/672232d010b0d582ee8c4905/Autumn_Budget_2024_web_accessible_.pdf.

11. The RT Hon Pat McFadden MP, “Reform of the state has to deliver for people,” December 9, 2024, <https://www.gov.uk/government/speeches/reform-of-the-state-has-to-deliver-for-the-people>.

there.¹² It's like we're driving a car with the brake stuck to the floor — no matter how hard we hit the accelerator, we still seem to crawl forward while destroying the engine. But it's not just about velocity. The steering column doesn't quite connect to the wheels; we try to head in one direction but end up off course. Then we try to head in multiple directions at once as a result of overlapping mandates. And to make things worse, we often hand off navigation to outsiders who don't understand the purpose of our journey. It's time to pay attention to how we're going to get to where we should be going. The machinery of government needs urgent work.

THE BRAKE IS STUCK

Government has a **pacing problem**.¹³ Societal needs and technological progress are changing rapidly, but the government is not. In fact, it falls further behind every year.



Government cannot keep pace with broader change because it creates so many forces that slow it down. You see this in metrics like the time to connect new power sources to our electrical grids. Despite the urgent need to modernize our energy infrastructure, the “average time projects spend in interconnection queues has grown substantially to five years, compared to three years in 2015 and less than two years in 2008,” according to a recent Niskanen Center report.¹⁵ Where speed is actually most needed, we are least likely to get it.

What are the forces making the government so slow? The first of those dysfunctions is what Nicholas Bagley of the University of Michigan calls the “procedure fetish,”¹⁶ and we dub the **bureaucratic anxiety cycle**. Anxiety about legitimacy and accountability drives critics to demand, and bureaucrats to seek refuge behind, more and more layers of procedure that show things have been done “by the book.” But all that procedure further erodes both legitimacy and accountability by overburdening the bureaucracy, reducing its ability to deliver meaningful outcomes.

12. Notable exceptions include Operation Warp Speed, Covidtests.gov, Direct File, etc.

13. Gary Marchant et al., *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem* (Springer, 2011).

14. POPVOX Foundation, <https://www.popvox.org/>.

15. Ted Boling et al., “Evidence-based recommendations for overcoming barriers to federal transmission permitting,” Niskanen Center, April 4, 2024, <https://www.niskanencenter.org/evidence-based-recommendations-for-overcoming-barriers-to-federal-transmission-permitting/>.

16. Nicholas Bagley, “The Procedure Fetish,” Niskanen Center, December 7, 2021, <https://www.niskanencenter.org/the-procedure-fetish/>.

The bureaucratic anxiety cycle



The examples of this trap are legion in government. Take the 5000 series procurement framework at the Department of Defense. Originally seven pages in length when it was written, it is now over 2,000 pages. That accumulation of procedure hasn't made defense procurement better, or faster. In fact, by its own admission, the Department's purchasing processes now take so long that the ships, planes, weapons, and software it buys are frequently obsolete by the time they are delivered.¹⁷ While process and procedure are a critical and normal part of any bureaucracy, we have allowed them to ossify and proliferate at a pathological rate.

This downward spiral is exacerbated by a complementary **anxiety cycle in the legislative branch**. Legislators are often frustrated with the performance of the administrative agencies they oversee. They frequently respond by being ever more detailed and specific in their instructions, reducing the agencies' sense of responsibility and accountability to deliver outcomes. Agencies are already highly constrained not just by the specific directions they've been given to execute in their area of focus, but by enormous bodies of rules governing how to conduct common operations from hiring to purchasing. The constant stream of mandates and constraints from above actually reduces the agency's effectiveness rather than improving it, furthering legislative anxiety about performance, and leading the cycle to repeat. While executive branch agencies must play their role in taming the procedure fetish, the first branch must do its part too.

The legislative anxiety cycle



17. Sandra Erwin, "Hyten blasts 'unbelievably' slow DoD bureaucracy as China advances space weapons," *Space News*, October 20, 2021, <https://spacenews.com/hyten-blasts-unbelievably-slow-dod-bureaucracy-as-china-advances-space-weapons/> and Ella Sherman, "The US military is 'alarmingly slow' at delivering new weapons, making it tougher to keep its edge against top threats, watchdog finds," *Business Insider*, June 18, 2024, <https://www.businessinsider.com/watchdog-dod-must-speed-up-weapons-deliveries-to-maintain-advantage-2024-6>.

Anxiety about accountability is often tied to concerns about safety or waste, fraud, and abuse. Many of these concerns are entirely reasonable, but policymakers often fail to understand both the weight and friction of the guardrails. It is reasonable to order an agency to study potential harmful effects of a new policy or technology, such as AI, but the existing accumulation of bureaucratic procedures mean such studies can take years to complete. Too often, the risk of inaction is not weighed against the risk of action. The accelerating gap between public and private sector capabilities has societal consequences. This larger context must be considered when assessing the effects of well-intended guardrails.

Another dynamic that causes policy accretion is what we call **the tyranny of tiny decisions**. Whenever a process is created in any large organization, there are always a series of suggestions for additional qualifiers and considerations that are individually proportionate and sensible responses to a possible risk. The problem is that when you look back and add up all the tiny, sensible things, you have created a clunking behemoth. And worse, because each individual addition makes sense, it is then incredibly hard to argue that it should be removed from the list, and given the work that's gone into it, also incredibly hard to wipe the slate clean and start again. Avoiding this outcome requires having a willing and able leader empowered to make choices about what's in and what's out, and dedicated to a usable, functional process, rather than one that makes every contributor happy.

The constant accretion of policy and process has another source: **adversarial legalism**. This, too, becomes a cycle. When government is sued or an agency called before a Congressional committee, it generally defends itself on the basis of having followed the proper procedure. But parties seeking advantage find in those procedures opportunities to attack or at least slow down government action. The attacks spur bureaucracies towards ever-more detailed, voluminous procedures, repeating the cycle. Both the lawsuits and the procedures are key drivers of the pacing problem. It is hard to move fast when you must generate mountains of paperwork and move at the speed of the courts.

With all these feet on the brake, it is not surprising that on some occasions, the government doesn't appear to move at all.

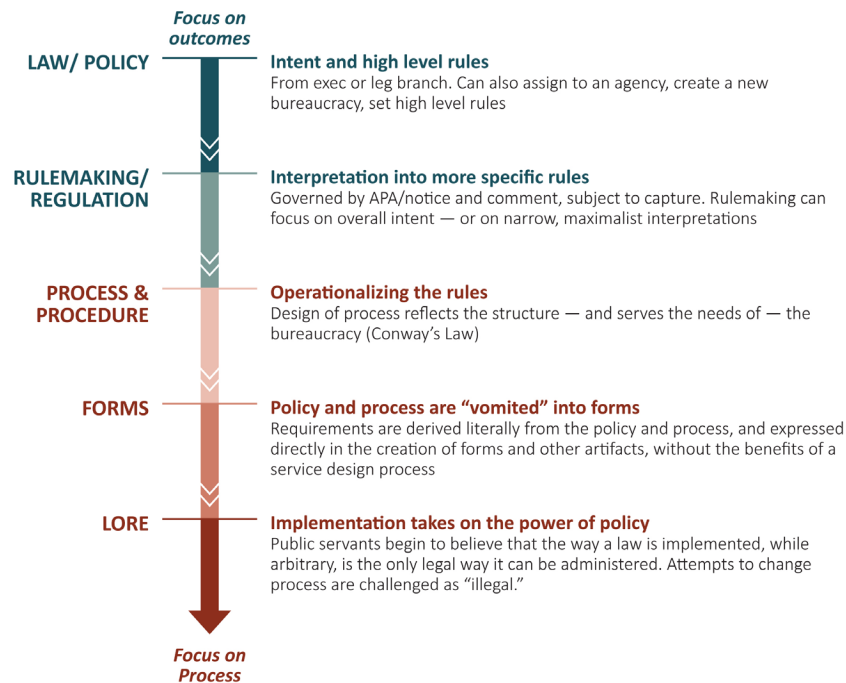
THE STEERING IS LOOSE

But it is not just the rate of change that's at issue. We increasingly see a gap between the intentions of law and policy and the outcomes they achieve in practice. If we unstick the brake, we must also reconnect the steering column to the wheels of the car.

One reason policymakers often misjudge the burden their guardrails will impose is a further dynamic we call the **cascade of rigidity**. The cascade begins with high-level principles outlined in legislation and executive orders. As these principles descend through layers of bureaucracy, they are translated into more specific and prescriptive guidance. Each level of government, from agencies and sub agencies to individual bureaus and divisions, interprets and operationalizes the guidance in its own way. With each step down the ladder, the flexibility intended by the original framers is diminished, replaced by rigid interpretations and narrow, literal applications of rules. Take, for example, how civil service rules intended to reduce bias and nepotism (and written at a high level to allow for judgment) have been translated at the operational level into a system that excludes the judgment of subject matter experts, and where instead candidates are evaluated on the basis of self-assessments and narrow, literal reviews by HR professionals. The result is that only those with insider knowledge of – or who are preferred by – the process can make it through to the stage where

hiring managers can consider them. The principles at the top and the mechanics at the bottom are in conflict. It can feel to lawmakers like the steering wheel they are supposed to have their hands on isn't properly connected to the wheels.

Cascade of Rigidity

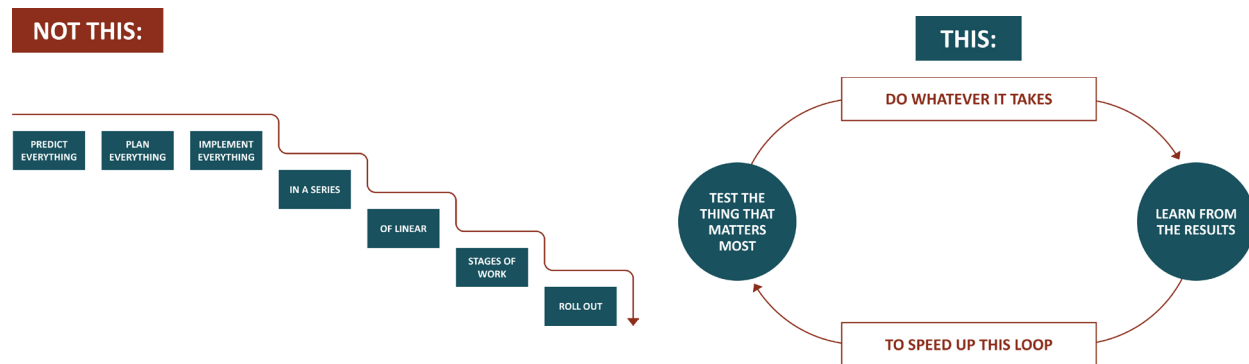


We also find ourselves somewhere other than where policy intended because we are trying to solve **complex problems with tools suited for merely complicated ones**. What's the difference? According to Theodore Kinni at the *MIT Sloan Management Review*, "Complicated problems can be hard to solve, but they are addressable with rules and recipes...They also can be resolved with systems and processes, like the hierarchical structure that most companies use to command and control employees."¹⁸ Many of the biggest challenges of the industrial age were complicated. It is complicated to build a road or a submarine, but many of the relevant factors are known: the force of pressure at great sea depths or the tensile strength of steel at a given temperature are well established. Waterfall processes (in which each of a series of consecutive stages like planning, development, rollout, and operations follows the next with no overlap) can work well when it's possible to predict the vast majority of the variables and gather all the necessary information up front.

By contrast, complex problems involve too many unknowns and dependent factors to reduce to rules and processes. If you look at the headlines we listed at the start of this report, most of them describe failures of government to address highly dynamic problems that feature many interrelated factors bound up in a web of massively interdependent systems, including procurement systems, legal systems, personnel systems, and systems of values and priorities. The laws of physics were not why these projects made headlines. It was many complex, rather than merely complicated, legal, bureaucratic, regulatory, and most importantly human factors that confounded even the best planning. And when we treat complex problems as if they were merely complicated, we are doomed not only to fail to solve them, but in many ways, to make them worse.

18. Theodore Kinni, "The Critical Difference Between Complex and Complicated," *MIT Sloan Management Review*, June 21, 2017, <https://sloanreview.mit.edu/article/the-critical-difference-between-complex-and-complicated/>.

Solving complex problems requires tight feedback loops — the ability to test, learn what works in the real world, and act on that knowledge. All the planning in the world can't guarantee a good outcome, not only because of the complex, interdependent nature of the problem, but also the fact that it is changing *while you are trying to solve it*. (Attempts at AI regulation are a glaring example here: the technology being regulated is changing faster than proposals for its regulation can even be considered.) The path to effectiveness involves iteration, testing, and learning in the field. But that is very hard to do when you **lack a feedback loop between policy and implementation**.



Today, policy implementation follows distinct phases that vary by circumstance, but largely share a one way flow from inception and design, often starting with the legislature, to implementation by executive branch agencies, and ultimately contact with the public, businesses, or other entities. Because the flow is one way, insights gained through this long process must often be ignored, because there is little leeway, on a practical basis at least, for returning to the authors of the policy with insights from the implementation and requests for adjustments. Implementation teams must follow the directives as written, even when it becomes clear that the outcomes are inconsistent with policy intent. This inflexibility is dangerous, especially in an ever faster-changing world.

It may be easier to envision tighter feedback loops if we look to examples abroad. In the U.S., the separation of powers between the executive and legislative branches creates structural distance between policy and its implementation. In the UK, by contrast, ministers in charge of departments are chosen from among the members of Parliament. (The equivalent in the U.S. would be if agencies were run by members of Congress.) This doesn't mean the legislative body has perfect insight into the challenges and opportunities of the bureaucracy, but under the right circumstances, it does provide for a more informed perspective.

The UK also makes provision for "secondary legislation," whereby government departments can draft and submit minor changes to laws within the overall framework of already passed legislation; in effect, to fill in the details. Parliament can either approve or reject secondary legislation, but it cannot amend it, so this tends to be a relatively quick process. This makes for a useful mechanism for making relatively minor tweaks to policy; approximately 3,500 of these amendments are passed in any given year.¹⁹ The UK Parliament's own website describes this route as a deliberate lever to put in place the "practical measures that enable the law to be enforced and operate in daily life."²⁰ Assuming ministers engage with the work of teams using test-and-learn approaches, they enjoy the option to adjust policy and legislation in near real-time, without

19. "What is Secondary Legislation," UK Parliament, accessed December 2024, <https://www.parliament.uk/about/how/laws/secondary-legislation/>.

20. *Ibid.*

having to go through the lengthy process of presenting a whole new bill to pass through the legislature. This sets them up well for taking advantage of what the team learns as they begin to implement, and adjusting accordingly. It provides a way to close the loop.

In the U.S. system, Congress may come to agencies' legislative affairs offices for technical assistance when a bill is being drafted, but even if they are candid, agencies have not begun implementation at that point — they don't know what they are about to learn as they engage with the real world problem, nor do they know the conflicts that will arise internally as they begin to operationalize the words of law. Even during implementation, agencies can struggle to communicate with Congress. As described in the sidebar on page 13, a U.S. Digital Service staffer working on implementation of a health care law broke protocol by speaking to the House Office of the Legislative Counsel on her own, because she knew her request for clarification would never make it through the relevant agency's legislative affairs process.

The U.S. system does not have a “secondary legislation” mechanism, though minor fixes to existing bills certainly occur through the regular legislative process. Appropriators often clarify or direct agencies through language in spending bills or in the accompanying reports. In the past, Congress also frequently passed “technical corrections” bills to address implementation issues, though these adjustments happen less frequently due to the governing-by-omnibus approach of recent years. Even if technical corrections occurred more frequently, they don't, under current practice, represent much of an opportunity for test-and-learn approaches, as they do in the UK. That would require the agencies themselves to employ test-and-learn strategies, which is often not the case. Without people in agencies learning what is working in the implementation and what isn't, there is very little that leadership and legislative affairs can proactively share with Congress even if the trust between the branches were far deeper than it is today.

These are big barriers, but the Supreme Court's recent *Loper-Bright* decision limiting courts' deference to agency interpretations of laws may erode some of the resistance to change. The system will need to adapt, and could be shaped towards greater agility. We should take inspiration from the glimmers of progress provided by our peer nations.

To gain the state capacity we need to address the challenges of the day, the executive and legislative branches must begin to operate together in “build-measure-learn” cycles. Only then will we feel that when our sometimes impossibly messy political processes actually work, and we take a hand to the steering wheel, we head in the direction we wanted to go.

WE'RE TRYING TO GO EVERYWHERE AT ONCE

Ezra Klein has critiqued the left for adding too many policy priorities into the design of programs, like the requirements that companies applying for funding under the CHIPS and Science Act detail their plans to address climate, gender, economic equity, and other considerations. Klein called this tendency “everything bagel liberalism,” and Matt Yglesias has called this “every policy objective, all the time, all at once.”²¹ To extend our car metaphor, it's hard to get somewhere in a timely manner if you decide you have to stop at 20 other destinations along the way. Quite often, you won't get there at all unless someone prepared to be unpopular chooses which stops get a visit – and which don't. Unfortunately, tradeoff denial is far too common.

21. Both are referencing the movie *Everything Everywhere All at Once*, in which a bagel with too many toppings becomes a sort of black hole.

CASE STUDY: OPEN LOOP ERRORS

MACRA (Medicare Access and CHIP Reauthorization Act) was designed to pay doctors more for higher-quality care. But an implementation team at the Centers for Medicare and Medicaid Services (CMS) knew that doctors were already frustrated with the burdensome and confusing ways they had to report their data under the existing program, and many were so concerned that the new system would be just as bad that they were threatening to stop taking Medicare patients. Thus, a law designed to improve the quality of care threatened to degrade it, especially for patients in rural areas who relied on the small practices that were most affected.

Recognizing how challenging the administrative requirements could be for practices with fewer resources and limited Medicare revenue, one provision in the law exempted doctors who treated a minimal number of Medicare patients. But CMS's initial interpretation of this provision would have required all providers to collect and submit a full year's worth of data in order to demonstrate they fell below the exemption threshold. This meant exempt doctors would still have to comply with all the program's requirements, including updating their systems and reporting data, only to be excused from all this at a later date. It's not hard to see why this approach, while technically accurate, would have worked against the intent of lawmakers. Those doctors would have left the program, hurting the very patients the law meant to help.

Another provision allowed smaller practices to form "virtual groups" to gain advantages enjoyed by larger practices. Staff interpreted this provision as a mandate to create a "Facebook for Doctors," a platform for doctors to find and connect with each other. A staffer on loan from the United States Digital Service, a part of the White House, doubted that Congress intended for CMS to create a social media platform, especially considering the limited time and resources available. She took the almost unheard of step of consulting the House Office of the Legislative Counsel, and confirmed that Congress simply wanted to make it easier for small practices to report together and had no intention of mandating a "Facebook for Doctors."

Under more common circumstances, these and other overly literal interpretations of the law would have resulted in a burdensome, unwieldy, and ultimately unsuccessful implementation. Doctors would have simply opted out, leaving patients with fewer options, and some in rural areas with none.

Thanks to nimble actions by people at CMS and USDS to ensure that Congressional intent was realized rather than over-relying on literal interpretations, this outcome was avoided. But conflicts like these all too rarely resolve in favor of common sense. Agency staff are commonly taught to treat legal language as literal operating instructions, as if a programmer had written code and they were the computer executing that code. But as any programmer will tell you, code rarely works as intended on the first try. It works after trying one approach, testing it, adjusting, and continuing that cycle over and over again. That cycle of adjustment is very difficult to engineer within policy implementation today.

We run on an open loop, in which implementation teams neither test their programs in the real world nor loop back to the source for adjustments. We need to build the affordances for them to do both, thus closing the loop. Otherwise, the code will more often than not run exactly as Congress wrote it, even if that doesn't result in what Congress wanted.

In 2017, for instance, San Francisco imposed a ban on doing business with states that failed to support LGBTQ rights and abortion rights, and even prohibited certain city employees from traveling to those states on city business, even to woo companies to relocate to San Francisco. The impact was disastrous. It increased paperwork enormously, in part because adhering to this policy was impractical and caused staff to apply for waivers to get around it — 538 in just one year — and in part because of the extra paperwork needed to demonstrate compliance with the ban for all purchasing. Far more staff time was needed to source new vendors for everything from computers to toilet paper. In one case, an LGBTQ-owned business that had been doing business with the city for years was cut from the supplier rolls when it was bought by a company from one of the 30 banned states. Prices went up, as suppliers who still qualified realized how much less competition

they now had.²² This law was cited as one of the reasons for a \$1.7M price tag for a public toilet²³ in a San Francisco neighborhood which became the focal point of constituent outrage. When it was repealed seven years later, not only had the ban failed to move any of the 30 banned states towards San Francisco’s preferred social policy, it had also had a disastrous effect on the city’s finances and operations.

In the technology world, there is a role designated to make those tradeoffs, and it’s telling that it has not until recently existed in government. That job is called a product manager. If you say “product manager” in policy circles, people hear “project manager,” a prolific title across government implementation teams. But the two are different. *Project* management is the art of getting things done. *Product* management is the art of deciding what to do.

Project management is critically important, especially in government where there are so many things to do. But that list of things to get done is often the superset of wishlists (at the project level, usually expressed as requirements) from a wide variety of stakeholders. If there’s no product manager doing the work of deciding what’s in or out, focused on the desired outcomes instead of checking hundreds of boxes, all the project management in the world will not result in a usable product. On the contrary, more project management means more bandwidth to add toppings to the bagel, until it becomes inedible.

This dysfunction is not limited to technology. It can exist at every level of the cascade from politics to policy to implementation of all kinds. One expression of it shows up frequently in procurement, as we saw with San Francisco’s ill-fated bans. The federal government also imposes a wide range of requirements on its vendors that are unrelated to the quality and cost of the goods or services being bought. And in many cases it should do so, as many of these requirements are meaningful safeguards against unethical business practices. But it is rarely one requirement, but rather the sum total of dozens or hundreds, that becomes a problem, as we described in the “tyranny of small decisions.” The authors of any given restriction aren’t thinking about all the ones their predecessors have imposed before, and aren’t practicing “product management” to ensure that by “adding features,” they aren’t making the product unusable. As Greg Jordan-Detamore explains, these rules raise prices by adding to the costs contractors incur, increasing their paperwork burden, and reducing competition when some vendors simply bow out.²⁴

Tradeoff denial is not the problem of one party, or of one Congress or presidential administration. Mandates and constraints come at agencies from both sides of the aisle, and they are cumulative. Very rarely are attempts made to go back and remove or revise what previous Congresses have imposed. Current leadership may care about their most recent initiatives, but that doesn’t make anything that’s been legally mandated in the past less mandatory. It is very hard to have priorities when nothing can ever fall off the laundry list of what needs to happen. If something is a true priority, it means that it will happen *at the expense of something else*.

WE’VE ABDICATED RESPONSIBILITY FOR NAVIGATING

At the risk of stretching our car metaphor too far, there is a final problem with how we are currently trying to achieve our policy goals. Government is ultimately responsible for getting the vehicle to its agreed-upon

22. Lyanne Melendez, “SF re-examines its procurement process after business ban on 30 states backfires,” *ABC News*, February 15, 2024, <https://abc7news.com/san-francisco-procurement-business-ban-boycott-sf-park-and-rec/14428958/>.

23. Heather Knight, “San Francisco Tried to Build a \$1.7 Million Toilet. It’s Still Not Done.” *The New York Times*, January 24, 2024, <https://www.nytimes.com/2024/01/24/us/san-francisco-toilet.html>.

24. Greg Jordan-Detamore, “Hidden Costs of ‘Everything Bagel’ Procurement,” *Civic Insider*, May 15, 2024, <https://civicsinsighter.com/p/hidden-costs-everything-bagel-procurement>.

destination, but we've outsourced too much of the work of navigating us there. While government will always need to purchase the fuel that propels the vehicle and hire out for maintenance, the directions that will get us to the destination need to come from within.

If you've worked in or with government, you've probably heard the term "one throat to choke." This refers to the need to have one prime contractor responsible for anything from the development of an IT system to the services a program is charged with delivering, and as the phrase suggests, to be able to blame them when things go wrong. The phrase is indicative of what has too often become a profound abdication of responsibility on the part of government. Well-intentioned attempts to ensure that government hires out for tasks that the private sector can and will do better have spiraled into a harmful ideology that assumes private contractors are better than government at pretty much everything and government should do as little as possible, of any kind of work, in-house. The reality is that if government is on the hook for the outcomes, which it should be if the work is funded by taxpayer dollars, public servants must make the critical calls, even if contractors may do some or even much of the work. And that requires that the competencies and capacity to make those calls — to navigate the vehicle to its destination — reside within government.

In 1966, the White House Office of Management and Budget released circular A-76, which built on previous policies stating that the federal government "will not start or carry on any commercial activity to provide a service or product for its own use if such service or product can be procured from private enterprise." The specific methods A-76 prescribed are no longer relevant; use of A-76 was prohibited first by the Defense Department and later across government when the press and public officials blamed shocking conditions at Walter Reed Medical Center on the process by which the Pentagon had competed and awarded the contracts to run the center. But the memo is important because it formalized the distinction between functions that are "commercial" and those that are "inherently governmental."²⁵ That dividing line between work that must be done by government employees and work that should always be outsourced has been hotly contested before and since the memo, and it desperately needs revisiting today.

This debate is too often framed as a binary: either you believe that government is bad at most functions and should outsource everything it can, or you believe outsourcing to the private sector (and the social sector, the abuses of which have gotten far more attention in recent years) is counter to public-sector values and want government staff to do it all. (Public sector unions of course have their own reasons for preferring insourcing, and it can be a surprise to leaders trying to, for instance, process a tenfold increase in unemployment insurance applications that hired help is not able to touch the claims because legal provisions make those the sole domain of what's called merit staff with many years of tenure.) But this is an unhelpfully reductive view of a complex issue. A more helpful debate would return to the core question of A-76 but return different answers, answers updated for the modern era.

For starters, the kinds of work that government contracts out for has changed. A-76 defined digital work broadly as commercial, and that thinking has pervaded government even as A-76 itself was rolled back. Digital is the domain probably most prone to the claim that government can never do it well. But an institution that doesn't understand digital is a very poor purchaser of digital products and services, as evidenced by the government's repeated failures in technology, from the Pentagon's struggles to even equip their staff

25. Congressional Research Service, "DOD A-76 Competitions," *In Focus*, December 22, 2016 <https://crsreports.congress.gov/product/pdf/IF/IF10566/3>. John R. Luckey, "OMB Circular A-76: Explanation and Discussion of the Recently Revised Federal Outsourcing Policy," Congressional Research Service, updated Sept. 10, 2003, <https://digital.library.unt.edu/ark:/67531/metacrs7717/m1/1/>.

with working computers²⁶ to the Department of Education’s debacle with the FAFSA financial-aid form this year²⁷ to the Commerce Department’s failed financial system that stopped tens of millions of dollars worth of invoices and reimbursements from being paid and shut down critical services in sub agencies such as the National Weather Service.²⁸²⁹

These failures are the result of a system that tries to buy software design and development services as if they were pencils or cars. It’s not just that we evaluate bids on inappropriate criteria, it’s that we’re not soliciting for the right deliverables. We lack the internal core competence to outsource well. And we can never fully outsource the development of the bespoke software that manages government’s unique missions, because the software is so intricately tied to operations. If public servants aren’t fundamentally in charge of what’s being built, they won’t understand how their departments function.

In the cases above, internal core competence would be pretty cheap. Proper leadership can often be established with a relatively small internal team, as we describe in Part 4. Given that contracts for development of government software frequently cost hundreds of millions of dollars, failing to invest in a modest internal team with the appropriate skills and authorities is extraordinarily penny-wise and pound-foolish.

The folly of skimping on internal core competence is well established in many domains. One paper by Zachary Liscow of Yale finds a remarkable correlation between the staffing levels of state departments of transportation and the per-mile construction cost of highways: increasing employment by one person per thousand in a population reduces costs by 26 percent.³⁰ When there is insufficient staff to manage these projects, their work is outsourced to consultants who often lack the continuity, context, or incentives to perform the work properly. There is ample evidence across a wide variety of government functions that contractors cost more than government staff for the same work, and that outsourcing at those higher rates causes the very people the government has trained and invested in to leave for higher-paying jobs in the private sector.³¹ In 2011, Former CIA Assistant Director Mark Lowenthal told Congress, “It’s the least experienced analytical staff since 1947, and this demographic trend will play out in years to come,” regarding the brain drain to the private sector.³²

Over-reliance on contracting exacerbates all the other dysfunctions we discussed above. In an environment of procedural bloat, contracting adds even more procedures. It adds even more layers in the hierarchy, meaning the cascade of rigidity has even more steps through which the intent of policy can be perverted. It adds even more temporal, organizational, cultural, and structural distance between policy and implemen-

26. Michael Kanaan, *LinkedIn* post, 2022, https://www.linkedin.com/posts/michaeljkanaan_technology-future-innovation-activity-6891726752759074816-2qCv/

27. “Botched FAFSA Rollout Leaves Uncertainty for Students Seeking Financial Aid for College,” GAO, September 24, 2024, <https://www.gao.gov/blog/botched-fafsa-rollout-leaves-uncertainty-students-seeking-financial-aid-college>.

28. Jack Fitzpatrick, “Commerce Delays Debut of Troubled Finance System at Chips Agency,” *Bloomberg Government*, April, 26, 2024, <https://news.bgov.com/bloomberg-government-news/commerce-delays-new-financial-system-at-chips-funding-agency>.

29. Jack Fitzpatrick, “Latest Disaster for National Weather Service: Paying Its Bills,” *Bloomberg Government*, April 8, 2024, <https://news.bgov.com/bloomberg-government-news/latest-disaster-for-national-weather-service-paying-its-bills>

30. Zachary Liscow, “State Capacity for Building Infrastructure,” Aspen Economic Strategy Group, December 4, 2024, <https://www.economicstrategygroup.org/publication/liscow-state-capacity/>.

31. Pratap Chatterjee, “Insourcing How Bringing Back Essential Federal Jobs Can Save Taxpayer Dollars and Improve Services,” Center for American Progress, March 28, 2012, <https://www.americanprogress.org/article/insourcing/>.

32. Mark Lowenthal, Testimony before the Subcommittee on Oversight of Government Management, the Federal Workforce, and the District of Columbia, “Intelligence Community Contractors: Are We Striking the Right Balance?” September 20, 2011, https://irp.fas.org/congress/2011_hr/contractors.html.

tation teams, making it far harder to engage in test-and-learn cycles. Software that's contracted out often can't be modified by government staff, making it hard to react when needs change. Top-performing teams at the Department of Veterans Affairs insist on having enough in-house staff on any given project that if a change needs to be made quickly, one of their engineers can do it. The norm, however, is to have to negotiate a change order or task order or, worse, begin a new procurement. This results in enormous cost and the loss of strategic flexibility and responsiveness. Government should not be beholden to suppliers, no matter how well known their brand or how global their reach, if it means they lose the ability to configure and adapt systems for constantly changing circumstances. If a government is locked into a multi-year contract and a technology we can't change, then the whole idea of test-and-learn is a non-starter.

Having internal competencies is sometimes seen as risky, because it's harder to argue you have that "one throat to choke." Vendors know that absorbing blame for failures is part of what they are being hired to do, and they don't mind because they know that they can (and will) continue to win contracts despite a history of failure, even very public failure, due to the nature of contracting rules. CGI Federal, famous for bungling the development of healthcare.gov, saw a miniscule drop in its valuation after it spent months in the headlines for the debacle. Since then, the stock price of its parent company CGI Inc. has nearly quadrupled, from roughly \$30 a share to \$114. Among the many government contracts the company has since won was a bid to work on the new FAFSA (albeit not as the prime contractor), a project that went as poorly as healthcare.gov.

The problems with FAFSA bring us back to our original question: what should be inherently governmental in today's world? There is a lot wrong with how we contract with the private sector, to be sure, but changing those procurement rules will only help so much if the core competencies government hires for and values in leadership are limited to policy-formulating expertise. The team at Federal Student Aid who oversaw the FAFSA rollout were not necessarily bad leaders or derelict in their duties. They may simply have been political appointees in a system that can only see the need for leaders with their particular skill set, leaving enormous gaps in agencies' capabilities. They came into FSA with a limited set of positions they could hire for, almost all defined for similar skill sets. The organizational structure they inherited, with all that entails (open billets, etc.) also assumed a policy-oriented leadership, with extremely limited affordances for operational leadership. Operational competencies were assumed to be contracted for, so the competency that should have been operational was instead represented by the procurement and contracting staff. Clearly, these procurement and contracting teams could have chosen vendors with more appropriate skill sets, as evidenced by the outcome, but contracting expertise is a poor substitute for in-house operational skills. We need a fundamental relook at the core competencies agencies hire for.

Rather than debating outsourcing v. insourcing on the basis of abstract theories about what is "inherently governmental," we should ask: what in-house competencies and capacities allow government to be able to outsource well? This, in turn, is critical to a larger question: how might government develop a far healthier contracting ecosystem, one in which companies win contracts because they are good at what they do, rather than good at playing a highly specialized contracting game?

Part 4: Finding our way

Improving the capacity of the American administrative state — fixing the machinery that serves us — is hard to prioritize and even harder to do. Previous administrations have not tackled it not only because of competing priorities, including a never-ending series of crises, but because the task understandably appears amorphous and insurmountable. Meanwhile, progress is hard to measure and therefore seldom tracked.

We address that amorphousness by breaking the problem down into four separate areas needing reform. The challenges described here are still monumental, to be sure. But they are not impossible. And if we are correct about the enormous benefit that an effective bureaucracy would deliver across policy domains, they are well worth doing, and even worth prioritizing over specific policy initiatives. Because no matter how politically beguiling it is to say, “Let’s just get some quick wins before all that tricky stuff,” the truth is, you don’t get many real wins unless you come to grips with the underlying obstacles.

The dysfunctions we identified in Part 3 don’t disappear when you name them, just as telling somebody they have cancer will not cure them. The dysfunctions must be attacked with a new set of tools, principles, incentives, and work routines. Here we describe the changes that will be needed in the bureaucracy of both the executive and legislative branches, and what will be needed from leadership for those changes to manifest and sustain.

If you want a government capable of achieving its policy goals — a government that can move quickly, break the bureaucratic anxiety cycle, manage complexity, and stay focused on priorities — you must:

1. Hire the right people and fire the wrong ones
2. Reduce procedural bloat
3. Invest in digital infrastructure, and
4. Close the feedback loop between policy and implementation

1. Hire the right people and fire the wrong ones

Government gets stuff done because of the people in it and the functions they are hired to fulfill. While the federal workforce includes countless dedicated, creative, and effective public servants, the system that selects, promotes, retains, and manages them is badly broken. Furthermore, too much of the work they are asked to do is the wrong work. Both of these issues need immediate attention.

REFORM THE CIVIL SERVICE

The foundations of the civil service are good. The Pendleton Act of 1883³³ was designed to move from a system of political patronage to one in which the majority of our government workers earn and keep their jobs based on merit and performance. The Civil Service Reform Act of 1978 established a set of Merit System

33. The Pendleton Act (enacted January 1833), <https://www.archives.gov/milestone-documents/pendleton-act>.

Principles³⁴ that describe a system in which “all employees...maintain high standards of integrity, conduct, and concern for the public interest,” “employees [are] retained on the basis of adequacy of their performance, inadequate performance [is] corrected, and employees [are] separated who cannot or will not improve their performance to meet required standards,” and “the Federal workforce [is] used efficiently and effectively.”

Principles are one thing. Practice is another. The chasm between the two is often huge, and glaring in practically every area of HR management.

Let’s start with hiring. In the federal government, 90 percent of hiring actions rely only on self-assessment questionnaires and an HR resume review that is frequently limited to scanning for the exact words used in the job description.³⁵ In other words, no independent effort is made to ensure the candidates have the skills they purport to have. In part because of that and other absurdities in the selection process as it is practiced today, only 51 percent of those job postings result in a job offer. In the other half of cases, hiring managers discard the entire process, which represents many months of work, rather than choose from the slate of candidates provided by their HR counterparts because it is evident they are not qualified. Among the postings that do result in a hire, that process takes significantly longer than it should.

The situation is arguably worse when it comes to managing out poor performers. According to a Merit Systems Protection Board study, “three-quarters of supervisors of unacceptable performers reported attempting ten or more different approaches for addressing the performance problem of their most recent poor performer.”³⁶ The same study concludes that even the “most effective” methods of resolving unacceptable performance available to federal managers are effective in “less than half of cases.” As with hiring challenges, this state of affairs affects productivity, retention, and motivation, as managers become consumed with HR processes while their low performers drag down the teams and their high performers become disillusioned.

In the vast space between hiring and firing, the picture is not much prettier. We could catalog similarly concerning statistics to illustrate problems in performance management, promotion criteria, job classifications, pay scales, and employee experience. Government may struggle to hire, but because of these many challenges throughout the employee lifecycle, we struggle equally with retention of high performers. The mission is so compelling it attracts top talent, but many quickly feel the grind of trying to drive with the brake stuck to the floor, and they burn out.

So why is there such a chasm between principle and practice? The gap lies in the host of bureaucratic policies and procedures that build up and around the original intent of the Merit System Principles. Think of

“If you have someone who’s not a good fit or incompetent or just not really interested in participating, there is not much you can do about it without dangerously neglecting the job you were hired to do. So it leaves me wondering what’s more important: HR’s processes or accomplishing the task the U.S. government has asked you to fulfill?”

— State Department employee

34. 5 U.S. Code § 2301.

35. “Hiring Assessment and Selection Outcome Dashboard,” Government Services Administration, <https://d2d.gsa.gov/report/hiring-assessment-and-selection-outcome-dashboard>.

36. “Remedying Unacceptable Employee Performance in the Federal Civil Service,” U.S. Merit Systems Protection Board, June 18, 2019, https://www.mspb.gov/studies/publications/Remedying_Unacceptable_Employee_Performance_in_the_Federal_Civil_Service.pdf.

them as barnacles on a boat, naturally occurring but in need of periodic removal. As administrators seek to apply the principles to ever-more detailed situations in routinized ways, they interpret and operationalize the principles in ways that result in outcomes inconsistent with the principles themselves. These derivative steps are subject to the “cascade of rigidity” described in Part 3, in which sound high-level principles descend through a risk-averse bureaucracy, becoming incrementally less flexible and more intrusively prescriptive at each level. The operational systems that rank-and-file public servants engage with operate vastly differently from what was intended by the authors of the laws and the principles they are meant to enact.

For example, one reason that most candidates for federal jobs are screened with the blunt tools of self-assessment and HR resume review before going to hiring managers is that the civil service has strict rules against bias and nepotism. The theory here was that concentrating hiring decisions among a small class of experts trained in avoiding these sins would prevent them from seeping in. But a major reason why we want to avoid bias and nepotism is that these tend to result in bad hires. Now, instead of hiring incompetent workers on the basis of prejudice, we do it on the basis of ignorance.³⁷

The hiring system needs to be renovated, but we don’t have to pull the whole house down. The foundations of the Merit System Principles³⁸ are still critically important for effective government delivery of services. To function as intended (and as desperately needed), the system needs to be stripped back to core principles and rebuilt properly, not thrown away at the expense of merit and fairness. This will require reenvisioning HR’s role to be more of a partner and less of a gatekeeper, moving more of the power to evaluate candidates and decide on the make-up of teams to the people responsible for the results those teams are charged with achieving. This will involve changes at all levels: statutory, regulatory, practical, and cultural. Our elected leaders must tackle the first two, and send clear direction to the bureaucracy on how to address the second two.

This is the threshold problem we must solve in order to make progress on the others. Without bold civil service reform, we will not have the state capacity to face the challenges of our decade and beyond.

REBALANCE THE BRAKE AND THE GAS

As important as getting the right people in the right positions is making sure to have the right positions in the first place. In particular, we need to redirect the government’s workforce from perfecting compliance toward pursuing results over the long term, and we must imagine what government jobs of the future will look like.

All systems must strike a balance between “go energy” and “stop energy,” similar to a car’s gas pedal and brake. Of course we need government agencies to exercise “stop authority” to ensure good governance, but our teams charged with achieving mission goals, with hitting the gas to meet critical national needs, almost always feel outgunned by those whose jobs are to hit the brakes. “We were six people trying to deliver the product,” one staffer told us. “There were easily 60 people telling us everything we couldn’t do.”

“We were six people trying to deliver the product. There were easily 60 people telling us everything we couldn’t do.”

37. Jennifer Pahlka, “AI Meets the Cascade of Rigidity,” The Digitalist Papers, September 2024, <https://www.digitalistpapers.com/essays/ai-meets-the-cascade-of-rigidity>.

38. “The Merit System Principles: Keys to Managing the Federal Workforce,” Merit Systems Protections Board, October 2020, https://www.mspb.gov/studies/studies/The_Merit_System_Principles_Keys_to_Managing_the_Federal_Workforce_1371890.pdf.

This is not simply a matter of cutting back on compliance-focused staff on a team. You have to simultaneously reduce the amount of compliance work there is to do. Otherwise you run the risk of making the already-damaging drag on the team's productivity even worse. Leadership must actively exercise judgment to decide what compliance is needed and how burdensome it needs to be, since, in the absence of direction from above, teams will default to the maximum of what might be required in order to cover all bases. We discuss this in greater detail in the next section.

That said, there is a case to be made that even when the compliance burden is taken as a given, staff resources are not being optimized to meet it. A major player in this question is the Office of Personnel Management. Despite the office's best efforts to help agencies achieve their workforce goals (and they deserve much credit for these efforts), agency staff still largely experience OPM as a compliance office, the cop who tells them that they can't hire the person they've worked for months to recruit and select because of a difference of opinion about whether they used the right hiring authority, for example. While agency staff are tasked with achieving meaningful goals, and are held accountable to Congress and the public for achieving those goals, OPM too often acts as a countervailing force that keeps them from securing the team they need to accomplish the task in time. This could be remedied in two ways: through delegation of authority from OPM to agencies, who should be trusted to appropriately interpret the law, policy, and guidance they are subject to, and through changes to that law, policy, and guidance to allow for greater use of judgment by teams to manage their workforces as needed to fulfill their missions. There could also be a role for AI in screening what are sometimes very high volumes of applications, provided there are team members who know how to use these technologies appropriately. Use of AI in screening could be enormously helpful if the HR teams are willing to take responsibility for good outcomes, not just fidelity to process.

But OPM's role in improving the government workforce goes far beyond some calibration around these compliance issues. More than just getting out of agencies' way to increase the velocity of the vehicle of government, OPM could and should take an active, strategic role in balancing the ratio of stop energy to go energy. Just as a company might look five years ahead at what roles it will most need to successfully compete in the marketplace, and plan to be able to fill those roles, so could OPM look at the needs of the federal government from a far more strategic perspective. As government works to right-size procedure through statutory, regulatory, practice, and culture changes, it must also actively plan for a workforce that can operate at the speed our nation needs it to. That will require a very different OPM than the one we have today.

2. Reduce procedural bloat

A police officer hired to patrol a beat to preserve public safety spends at least 80 percent of his time on paperwork.³⁹ A technologist hired to develop a web application for a state agency spends more time on compliance documentation than writing code. A procurement officer for the Defense Department spends shockingly little time actually evaluating and purchasing goods and services to meet the needs of our national security. While our first pillar highlights the need for government to hire the right people for the right job, aggressively reducing the burdens that keep our existing workforce from achieving the outcomes taxpayers believe they are paying for could go even further towards increasing state capacity, if done alongside wider reform.

39. Gagan Biyani, "Why does it take hours for Oakland police to show up?" *Empower Oakland*, August 9, 2024, <https://empoweroakland.substack.com/p/why-does-it-take-hours-for-the-police>.

TAKE ON POLICY, PROCEDURE, AND REGULATORY ACCRETION

As suggested by our discussion of the “procedure fetish” in Part 3, “right-sizing” procedures to be fit for purpose would go a long way toward unsticking the permanent brake in the metaphorical car.

Congressional and executive leaders should engage in large-scale inquiry into the appropriateness of the procedures mandated by law and policy by asking three questions:

1. What harms is a procedure intended to protect against?
2. Are those harms well addressed by the procedure as practiced today?
3. Is the protection against those harms, as practiced today, worth the costs of imposing the procedure?

Principles for de-proceduralization

A complete and detailed agenda for de-proceduralizing any given domain is out of scope for this report. However, we offer core principles to shape future efforts:

- The revised system must shift its emphasis from compliance to meeting mission needs. This means power to make decisions must shift from compliance personnel to the people closest to the work.
- The system must be able to work much faster than it does today. Speed of action will be an important metric in every area under reform.
- In some cases, tweaks to the system are no longer sufficient to change on-the-ground behavior. Change must be big and structural. At the same time, change must be grounded in the public interest and uphold such core principles as expressed in the Merit System Principles (in the case of HR reform), anti-corruption principles (in the case of procurement), etc. In seeking much-needed bold change, we can’t throw the baby out with the bathwater.
- We must accept that judgment is an essential quality of the work of public servants. That means decisions (hiring, firing, learning from user testing, etc.) should be defended in the first instance for being substantively reasonable, not for being made with pedantic fidelity to process.
- Change must be both top-down and bottom-up, but to properly diagnose dysfunction and avoid unintended consequences, the bottom-up view is critical to understand. Look at the artifacts of the day-to-day-work — forms, templates, documentation — not just policy as written, to understand and affect how that policy has been operationalized.
- Where possible, move processes from a default of no to a default of yes, with affordances for catching and fixing bad results.
- Accept that complexity itself is the problem. This simple point has profound implications. For example:
 - Adding new exceptions and carve-outs to speed up a given task increases complexity, further undermining the overall system. Fix the overarching procedures instead.
 - Eliminate the proliferations of pathways and authorities. As new tasks arise, existing mandates should be broadened, not buried under new layers of authority.
 - Delegate authority and responsibility to the agencies rather than micromanaging them by statute.

Right-size procedures that impact every agency

The vast majority of procedures are bespoke, built to meet a specific need within a specific agency. These can only be addressed by wide-ranging inquiry by agency leadership and the Congressional committees who oversee them using the rubric described above. But there are also procedures that affect all agencies, all policy domains. Every agency is affected by the procedural burdens of compliance with human resources policy, procurement policy, and cross-cutting laws like the Paperwork Reduction Act (PRA). Efforts to right-size these procedural burdens would impact the real productivity of some significant number of the 3 million federal government employees.

The PRA was written in 1980, a decade and a half before the commercial Internet. It prescribes a classic waterfall process for approving the collection of information by the federal government — a process that now averages nine months and often runs much longer. The goals of the PRA are better served through the adoption of practices of the digital era, such as user research, usability testing, and instrumentation of websites that allow for an ongoing understanding of the experience of users. Recent guidance from the Office of Information and Regulatory Affairs (OIRA) improved the implementation of this law, but statutory reform remains necessary to reduce unnecessary paperwork and incentivize agencies to update their practices for the digital era.⁴⁰

Procurement policy provides another example of the value of the rubric we established above. Reformers were right to enact procedures to protect against corruption, waste, and favoritism, particularly in how government bought the goods and services it needed. A return to a system without guardrails would be unwelcome and untenable. But those guardrails have grown in complexity over time, and they are so costly as to serve as their own barriers to competition, causing government to rely on contractors who excel at navigating the procurement system, not at delivering the best solutions. The original harms are not well addressed by the current procedures, nor are the costs worth the benefit they seek to provide. Add to these challenges the tendency to burden procurement processes with unrelated policy goals, as we discussed in Part 3. As with our HR framework, the principles remain sound, but decades of incrementally greater detail have turned the system as it operates today against those principles. Right-sizing the policy and procedure framework to strip back much of its accumulated complexity would better serve the goals of the system, and increase the ability of government to achieve its policy goals.

Reduce policy accretion

At the root of these out-of-control procedures is out-of-control policy and regulatory growth. Pentagon policy suffers from this mightily, weighing in at the equivalent to 100 copies of “War and Peace.” This has consequences not just in the bureaucracy, but on the battlefield. Warfighters are required to assess planned operations’ compliance with law and policy before acting. As one former DoD employee put it, “If the time to understand the policies exceeds the time window, the operation gets canceled.”⁴¹ There are plenty of reasons an operation should be canceled, but failure to understand our own rules should not be one of them.

40. Ben Bain and Jennifer Pahlka, “A win for good government and commonsense: New guidance on the Paperwork Reduction Act,” Niskanen Center, November 26, 2024, <https://www.niskanencenter.org/a-win-for-good-government-and-commonsense-new-guidance-on-the-paperwork-reduction-act/>.

41. Jennifer Pahlka, “Maligned Pentagon AI program ‘GAMECHANGER’ might inspire Congress to change its own game,” The Hill, October 12, 2023, <https://thehill.com/opinion/congress-blog/4253368-maligned-pentagon-ai-program-gamechanger-might-inspire-congress-to-change-its-own-game/>.

And this example barely scratches the surface of the challenges this policy cruft⁴² can create for operational effectiveness, in the Pentagon and far beyond.

Policy change is political. There are vested interests who will resist many efforts to simplify policy, for example in the case of contracting. But there is a place to start that's much more neutral and objectively administrative. We have hit these untenable volumes of policy accretion at a time when new tools are available for tackling them in the form of generative AI. No LLM can resolve differences between disagreeing parties, but it can help identify provisions in policy and regulation that are vestigial or de minimis, or conflict with other regulations that are equally low-stakes.

This is starting to happen. A team in the Ohio's Lieutenant Governor's office used AI to discover 2 million words that could readily be removed from the state code.⁴³ The Department of Health and Human Services similarly undertook a Regulatory Clean-Up Initiative at the end of 2020.⁴⁴ The Army, burdened by decades of accumulating requirements governing the acquisition of equipment, has identified and inactivated 52 legacy requirements documents that have been holding the service back from prioritizing "war-winning capability" over merely "additive capability."⁴⁵ The war-winning capabilities are the right toppings on the metaphorical bagel. The rest were weighing down our warfighting platforms and diluting our military effectiveness.

REDUCE THE SURFACE AREA FOR LEGAL ATTACK

Public servants tend to focus on process and procedure in part because their agencies are so frequently sued. Proper process and procedure help — but do not always succeed — in limiting the number and impact of those lawsuits. The last century has seen a mostly steady increase in the rights of individuals and interest groups to sue the state, on the theory that these rights would constrain the worst impulses of government and make it more responsive to the public's needs. While the increase in those rights has had many positive impacts, the time has come to consider whether the right to sue has so effectively crippled the state into inaction and tied it into bureaucratic knots that the needs of the public are not in fact being met.

This issue is central to several other problems we have already discussed, including the health of the civil service. Disgruntled employees sue their employers regardless of sector, but the public sector provides them far broader grounds for doing so, and for preventing or postponing actions such as employment termination. Supervisors of poor performers who try to fire them can easily find significant portions of their working time consumed with avoiding or handling lawsuits. (Many supervisors just work around their problem employees.) Some of these, of course, are substantive claims of wrongful treatment. Many of them, however, are frivolous, and the law, policy, and regulations governing this domain must do a better job distinguishing between the two.

It's important to understand the nature of many of these lawsuits, whether in federal civil service, permitting energy infrastructure, or elsewhere. Current statutes allow for many kinds of procedural objections. That means that an employee doesn't necessarily need to prove wrongful termination, or an environmen-

42. "The outdated, redundant, overly-complicated, or conflicting elements within our laws and regulations." Ben Bain and Christine Tsang, "Cutting Through 'Policy Cruft,'" Niskanen Center, October 30, 2024, <https://www.niskanencenter.org/cutting-through-policy-cruft/>.

43. Ned Oliver, "Ohio uses AI to eliminate unnecessary words in state administrative code," *Axios Columbus*, April 29, 2024, <https://www.axios.com/local/columbus/2024/04/29/artificial-intelligence-ai-ohio-state-administrative-code-husted?stream=top>.

44. Department of Health and Human Services, "Regulatory Cleanup Initiative," *Federal Register* 85, no. 221, page 72899 (November 16, 2020), <https://www.govinfo.gov/content/pkg/FR-2020-11-16/pdf/2020-21774.pdf>.

45. Jen Judson, "Officials move to purge stale weapon requirements from Army books," *Defense News*, October 14, 2024, <https://www.defense-news.com/land/2024/10/14/officials-move-to-purge-stale-weapon-requirements-from-army-books/>.

tal group harm to the environment, in order to file suit. All each needs to do is assert fault — often minor fault — in the procedure that led to the decision being challenged. That procedure doesn't even need to be required by law; complaints can assert, for example, that because previous environmental studies included a certain element, then a study that doesn't include that element, even if it is irrelevant to the circumstances, is incomplete and therefore legally deficient. Those complaints may not be taken seriously in court, but fear of them nonetheless drives bureaucratic behavior. Environmental impact statements now take an average of more than four years to complete and average 600 pages⁴⁶ in part because agencies are trying to protect against lawsuits, substantive or frivolous. Combine this legal indulgence with the accumulation and accretion of procedure in virtually all government functions, and the overly broad surface area for attack becomes apparent.

Fixes will vary by domain, and we have already discussed reforming civil service rules to limit the opportunities for litigation to substantive and meaningful wrongdoing. Some versions of permitting reform recently proposed contain elements along these lines as well. But a comprehensive agenda to reduce the risk aversion created by frivolous lawsuits would involve reforms to the Administrative Procedure Act, the 1946 law that established some baseline procedures for how administrative agencies must propose and establish regulations. The courts have interpreted the law expansively, subjecting virtually every agency action to close judicial scrutiny—the proverbial “hard look.” To guard against the hard-to-gauge risk of judicial invalidation, agencies devote immense amounts of time and energy to bulletproofing administrative records instead of doing their jobs. At a moment when we desperately need to build everything from transmission lines to freeways to housing, it's time to update the basic framework of administrative law to fit the country's needs.

RESIST TRADE-OFF DENIAL

In Part 3, we discussed the problems of laws, policies, and practices that try to achieve too many goals at once, thereby achieving none of them well. Ultimately this is an issue of setting priorities and operating to those priorities. As we said earlier, if something is a true priority, it means that it will likely happen *at the expense of something else*. Our leaders in the executive and legislative branches are not always going to agree on priorities, but where there is agreement, there must also be a recognition of what can give way to make room for what now needs to be done.

Clearer priorities start with policymakers understanding that they must subtract as much as they add — for every new mandate or constraint they impose upon an agency, they should find one to remove. When mandates pile up, there is simply more to be done than the agency can do well; when constraints pile up, the agency is crippled by rules and its leaders find themselves playing Mother May I with Congress and other oversight bodies, unable to take a step forward without permission.

CASE STUDY: FINDING THINGS THAT CAN JUST GO

There are some well-intentioned programs that just don't achieve much. The Social Security Administration, for instance, still offers a “Lump Sum Death Payment” of \$225 to surviving spouses or children of eligible workers whose jobs were insured by Social Security. Originally designed to help families with funeral costs, the benefit program was established in 1935. Congress capped the amount in 1954 and has not revisited the issue since. Today, the median cost for a funeral is almost \$8,000. While there are clearly families who could benefit from \$225 at a difficult time, few families persist through the administrative burden of applying for a one-time check that is relatively so small to actual costs. No matter how few families apply, the costs of administration and reporting on any program are not negligible. It's not clear that the juice is worth the squeeze, as they say.

46. “FACT SHEET: CEQ REPORT ON ENVIRONMENTAL IMPACT STATEMENT TIMELINES (2010 – 2018),” Executive Office of the President Council on Environmental Quality, June 2020, https://ceq.doe.gov/docs/nepa-practice/CEQ_EIS_Timeline_Fact_Sheet_2020-6-12.pdf.

This is a tough prescription. On the left in particular, elected leaders run on what they'll do, not what they'll undo. Undoing is often boring, detailed work (though today AI can help). And even if the goal is generally deregulatory, it's easier to craft a new bill with a new rule or mandate than to go troubleshoot existing statute or regulation. In the mid-2010s, there was a slew of scandals about EPA employees watching porn on government computers during working hours, and the difficulties in firing them. One employee admitted to watching two to six hours per day of pornography, and four months after having been discovered was still employed by the agency. Since existing policy already prohibited using a government computer for any unauthorized purpose, and should have allowed for dismissal over the lack of productivity even without the inappropriate nature of the material, one response to this incident could have been to deeply investigate how the Merit System Principles and other relevant policies had been operationalized so poorly as to allow for this outcome. In other words, fix what's wrong with the implementation of the existing law. But at least one member of Congress, a Republican, introduced a clearly unnecessary bill to ban the use of pornography on government computers.⁴⁷ The bill didn't pass, but it's an example of the temptation on both sides of the aisle to layer on instead of excavate and renovate.

Undoing is also hard because what's being undone may have champions who will actively resist, sometimes because they benefit from what you're undoing, sometimes simply because change is hard for people and organizations. But policymakers must understand that while the words of law they write are meant to be magic, that magic doesn't work if the machinery of implementation is congested.

3. Invest in digital infrastructure

Government's digital capacity and competencies are downstream of its ability to organize around outcomes, to reduce procedural bloat, and to hire and retain the needed workforce. So far we've discussed these bigger issues because you can't get technology right without addressing them. But almost all implementation in government relies on, is enabled by, and feeds into digital and data infrastructure. And getting the tech right, and adopting tech industry practices, can help organize around outcomes.

TRANSITION TO PRODUCT MODEL FUNDING

Investment implies higher spending. And we do need to spend on platforms and teams, both discussed below. But within the frameworks dominant in government today, higher spending in digital rarely results in better software. We are often spending too much, and spending it badly. If we want to build resilient, scalable software that works at a cost roughly commensurate with the private sector, we must change the model that governs how we do it.

How bespoke software is built is currently dictated by how it is funded. By funding custom software as a discrete thing to be bought instead of a capability to be developed and constantly updated, we make it both worse and more expensive. (Commodity software, like back-office systems that handle payroll and accounting or productivity suites like Microsoft's or Google's, is a different story.) This is best illustrated through a series of graphs, each one entirely fictional but representative of two fundamentally different approaches to funding, and thus development.⁴⁸

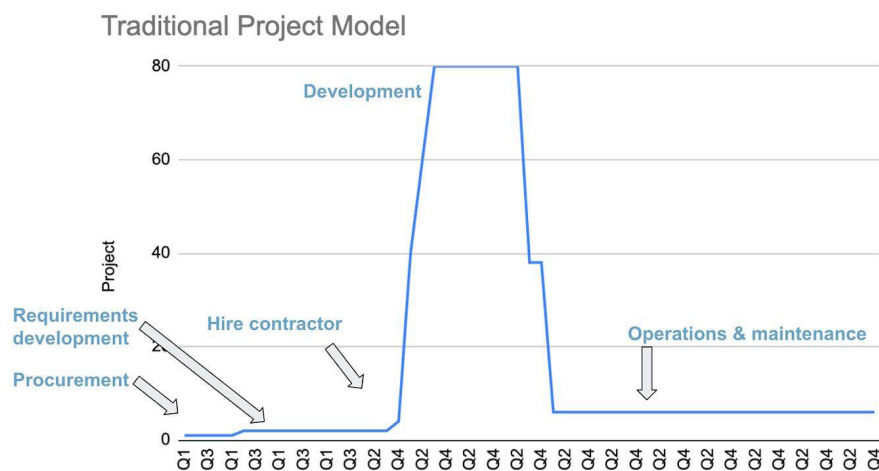
47. Colby Itkowitz, "Congressman seeks ban to stop federal employees from watching porn all day," *The Washington Post*, Sept. 14, 2014, <https://www.washingtonpost.com/blogs/in-the-loop/wp/2014/09/24/congressman-seeks-ban-to-stop-federal-employees-from-watching-porn-all-day>.

48. This section is adapted from Jennifer Pahlka, "Project vs Product Funding," *Eating Policy*, June 3, 2024, <https://www.eatingpolicy.com/p/project-vs-product-funding>.

Government typically follows a “project” model. The following graph shows the number of staff who work on an IT project at its outset, as requirements are being developed, a request for proposal written, bids from contractors sourced and evaluated, and a winner chosen. The contractor, once hired, brings on a team to develop the software based on the request for proposal (RFP), and the staffing levels (counting both internal and contracting staff) shoot up. There is a development period, followed by a short period of “user acceptance testing,” and then the project falls into “operations and maintenance” (or O&M), which is a different line in the budget.

Contrast this with a typical “product” model, depicted in the following graph, in which instead of a requirements-gathering phase up front, a small team (often but not always internal to government) conducts what are called discovery sprints to better understand the problems the software is supposed to address. If some parts of the proposed solution are riskier than others (for instance, it’s not clear whether a data integration will work well), they find ways to test those problems first, before an entire software solution has been built. They may develop prototypes to help question their assumptions, and they engage with the users of the software (often multiple sets of stakeholders) from the beginning. Product teams almost always leverage contractors, but the contractors are there to complement a core internal team which holds the product vision and provides clear direction to vendors. Staff is added slowly over time as the team learns what they need, but doesn’t dramatically ramp down once a first or even second version is shipped. As Dave Guarino quips, “Google didn’t lay everyone off after they stood up search.” Indeed, they invested more.

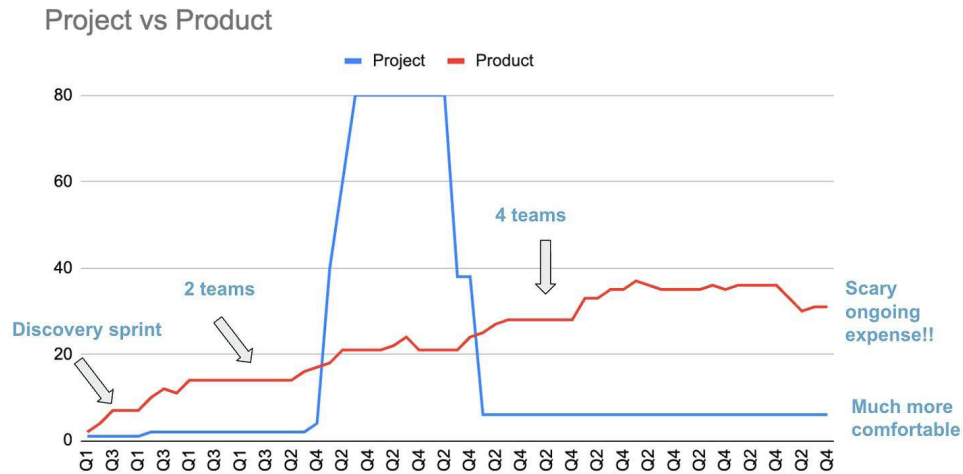
At this timescale, there seems to be an obvious reason to prefer the project model: the minimal ongoing



expense. This is what appropriators and oversight staff in Congress and elsewhere seem to be looking for as a sign of success. But as anyone following government’s challenges with technology knows, this is not the right timescale to look at.

What happens next on the timeline is one or more of the following scenarios: the software doesn’t work well for its users, and funds are sought to fix its defects; it quickly becomes outdated as a result of changes in the technical environment, the policy environment, or other external factors, and funds are sought for modernization; or new needs have emerged that the existing software doesn’t address...and funds are sought to meet those needs.

In the longer term, the project model usually shows more than one spike, often an ongoing series, as mod-



ernizations fail, needs escalate, and even more money is allocated. We end up building concrete boats, things we know don't make sense, but the incentive structure makes it so people do what they are told at the expense of successful outcomes. Put into contrast with the slow and steady product model, the project model becomes much less attractive on a purely cost basis.

But cost is far from the only reason to prefer the product model. The project model assumes that development work is at some point "done," but that is not the case. If software isn't updated, it rots.

“Software is never done”

- Software Acquisition Practices, Defense Innovation Board⁴⁹

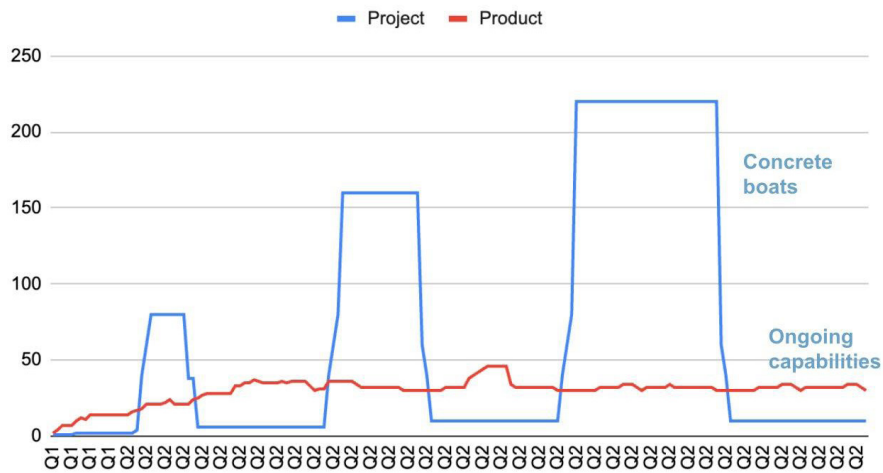
The product model is not only less expensive in the long run, it results in working software that doesn't need "modernization" of the kind you've become used to hearing about because it's constantly being updated and improved. To extend Dave's quip, you don't hear about Google pausing search for a modernization. Google Search first came out in 1998, and it is unlikely that its code base today bears much relation to what existed then. But Search is a product, not a project. It never went into O&M. It changed daily, not every other decade.

Putting this into practice is transformative. New Jersey's unemployment insurance team, for example, made improvements every few days throughout and beyond the pandemic.⁵⁰ The federal Department of Labor crowed about that state's successful new system, a success built on "a continuous approach to IT modernization over an all-or-nothing strategy."⁵¹ Dave Cole, New Jersey's Chief Innovation Officer, said: "Far too often,

49. "Software Acquisition and Practices (SWAP) Study," Defense Innovation Board, May 2019, <https://innovation.defense.gov/software/>.

50. Jennifer Pahlka, "Better government tech starts with people. New Jersey shows how." *The Washington Post*, June 13, 2023, <https://www.washingtonpost.com/opinions/2023/06/13/new-jersey-digital-unemployment-insurance/>.

51. "U.S. Department of Labor and New Jersey Highlight the Successful Partnership to Streamline the Delivery of Unemployment Insurance," U.S. Department of Labor, May 2024, <https://www.dol.gov/agencies/eta/ui-modernization/nj-press-conference>.



other states, large and small, have spent hundreds of millions of dollars to do one monolithic overhaul of their UI technology and applications, only for the resulting experience to remain just as confusing, just as frustrating, and just as demoralizing for claimants and state UI staff.”⁵² It’s great to see a state escape that trap.

As much as it’s tempting to take agencies to task for their addiction to the project model, this is not something they can fix on their own. It should be easy for agency leaders to assign small teams to work on discovery, test assumptions, and even build prototypes to help them learn about the problems that need solving and de-risk future development. But justifying technology development of any sort requires incredibly detailed business cases. The very fact that business cases take so much time and effort to construct disincentivizes teams from asking for small amounts of money to start small and test assumptions. Why ask for \$1 million to get started responsibly when asking for \$100 million would take just as long, and get just as much scrutiny? Many agency leaders will tell you it’s actually harder to get funding for \$1 million than \$100 million.

Government has many distinct needs, and private sector parity should not always be the sole goal. But it is telling that the companies with the best track records creating value through digital delivery employ the product model. And those companies are funded in ways that complement that model. In particular, venture capital firms fund startups to launch small and learn quickly, then invest more when the products get traction and show value, and the same pattern is used within large enterprises as well. Incremental, value-based funding pairs perfectly with small, cross-functional teams that practice agile development and focus on understanding and meeting user needs. On the other hand, spending a decade preparing for a project, then approving \$600 million to spend the next decade building it, pairs perfectly with giant waterfalls in which enormous teams check meaningless boxes to deliver on requirements that are many years out of date.

To make this shift, Congress would need to enable ongoing funding streams that lead agencies to develop in the product model. In tandem, Congress should work with OMB and agencies to change the laws, regulations, processes, and practices that impede agencies from operating in the product model.

We noted at the start of this section that this recommendation does not apply to the way government purchases commodity software. Those practices are also in need of reform to stop the proliferation of unnecessary requirements that lead agencies to build what they could much more effectively and efficiently buy. The next administration could find significant cost savings by addressing these and other dysfunctions of

52. *Ibid.*

technology purchasing as well, but the real wins are in getting the software that runs our most critical missions to work better. As Jennifer said in her book *Recoding America*:

*Though government should buy commodity products for commodity functions, when it's not accounting or payroll but your agency's mission, the technology needs to be your product. It can't just be a project that was contracted for, developed, tested, and declared "done." You need to own the code, and you need to be able to change it to meet your needs. This doesn't mean that you can't use contractors at all—in government, you will almost certainly use them. It means that you must have the core competencies to support a living, ever-adapting system. Government knows how to acquire technology. What we need to acquire are capabilities.*⁵³

INVEST IN CENTRALIZED AND DECENTRALIZED TEAMS

The required investments in staff to move from a project model to a product model are small compared to the costs of outsourced mega-projects, especially when those fail. Similarly, creating centralized digital capacity in offices like the U.S. Digital Service (USDS) and the Technology Transformation Service (TTS) generates enormous returns on investment. These teams not only intervene when projects go off the rails, but they spread the product model and other best practices throughout the federal government, preventing future disasters and building up capabilities within agencies. To name just one example, they've helped the Social Security Administration save \$280 million in infrastructure costs and reduce their call center wait time from over 42 minutes to under 12 over the last quarter.⁵⁴

Despite their visible successes, recent budget cycles have not been kind to these units. USDS used to have its own appropriations but is now funded by a 'pass the hat' mechanism whereby the agencies they work

CASE STUDY: ONLINE PASSPORT RENEWAL

When a team from the Bureau of Consular Affairs at the Department of State reached out to the U.S. Digital Service, they were well down the road on their second attempt to build a system to allow passport holders to renew online. The first attempt had gone badly, eventually being shut down due to performance and data integrity problems. The contract had been scrapped and the effort restarted. The rebooted project was doing better, but was still not benefiting from best practices of consumer software development.

What the USDS team found was functional code, but a lack of user research or usability testing that would make the system simple for applicants. Without proper instrumentation, the team wouldn't know what was working and what wasn't. It was also hard for the Consular Affairs team to make changes to their system. Whereas modern software teams should practice what's known as continuous integration and continuous delivery (CI/CD), the Consular Affairs team was limited by a platform and processes that only allowed for changes every two months. It's hard to improve your software based on tests and feedback from users if you can only change it so infrequently.

The USDS team helped Consular Affairs procure a new platform that allowed for CI/CD, and introduced the team to the practices they would need to take advantage of it. They showed them how to set up usability testing to better understand what applicants would struggle with and what to do about it. In the course of helping to launch the first-ever online passport renewal, they also whetted their partner's appetite for modern tools and processes. The Consular Affairs team could now see how much better online passport renewal could be using these tools, and set about creating a new version that's even better, which will launch in early 2025. And other user-facing services, like visa applications, could soon take advantage of the scalability and flexibility the team now values. It started with support on one project, but USDS has now been the spark for widespread transformation in an agency that Americans rely on.

53. Jennifer Pahlka, *Recoding America* (New York: Metropolitan Books, 2023).

54. "Increasing customer satisfaction on SSA.gov by 53%, U.S. Digital Service, 2024, <https://www.usds.gov/impact-report/2024/ssa/> and "800 number performance," U.S. Social Security Administration, 2024, <https://www.ssa.gov/securitystat/800-number-performance>.

with chip in or directly cover the cost of the staff assigned to them, turning empowered change agents into something more like consultants who must please the client. The Technology Modernization Fund (TMF), an innovative funding vehicle to accelerate federal IT modernization, received \$1 billion in the American Rescue Plan Act, but Congress and the Biden Administration have failed to provide sustained funding. We are going in the wrong direction when it comes to the digital infrastructure we need, both technical and human. Should the new administration and Congress make these centralized teams and funds a priority, it will pay off in multiples of hard spending reductions elsewhere, while setting the stage for agencies to build their own decentralized capacity with new teams of their own.

INVEST IN PLATFORMS

When you pull apart government services into their constituent parts, lots of the individual components are fundamentally the same. Filling in a form, finding out whether you are eligible for a program, taking a payment, sending a query about the status of your application, verifying that you are the person you claim to be when using a service — all of these elements apply equally to a huge range of interactions between citizens and the government.

Government organizations tend not to see it that way. Applying for a driver's license and applying for a visa might involve very similar component parts, with different rules on top. But the teams running them will generally build, buy, and run every single part of each service specifically for them. And all across the government, thousands of services will be doing the same — repeatedly rebuilding very similar components again and again, at great expense and effort. It's like building a whole kitchen every time you want to have a meal.

There are various reasons why this keeps happening. The way accountability runs in government means that leaders tend to be wary about creating dependencies on other parts of the state to run pieces of their service, where they are the ones who will be shouted at even if the piece that goes wrong isn't in their direct control. However, this desire for control is cultural as much as it is rational. Everyone likes to think their service is special and unique, even when this is clearly not the case. Fundamentally, the losers in this arrangement are taxpayers, who pay handsomely for all this unnecessary repetition, and users of government services, who have to navigate a different experience every time they interact with a different piece of the public sector.

Many governments around the world have now begun to build common digital platforms — component parts of services that can be iteratively improved and reused repeatedly across many different areas of government.

- In the last decade, the UK government developed a series of shared technology platforms like GOV.UK Pay⁵⁵ (to receive payments) and GOV.UK Notify⁵⁶ (to send updates to citizens on application status).
- Estonia developed the X-Road, a distributed information exchange platform that makes it possible for different systems to communicate all across the governmental sector.⁵⁷
- India's Aadhaar system, described as the “the most sophisticated ID program in the world,”⁵⁸ is

55. GOV.UK Pay, <https://www.payments.service.gov.uk/>.

56. GOV.UK Notify, <https://www.notifications.service.gov.uk/>.

57. “X-Road - interoperability services,” e-Estonia, last accessed December 2024, <https://e-estonia.com/solutions/x-road-interoperability-services/x-road/>.

58. “‘Aadhaar’ most sophisticated ID programme in the world: World Bank,” Daijiworld, March 16, 2017, <https://www.daijiworld.com/news/newsDisplay.aspx?newsID=442948>.

relied on by public and private sector services alike, with over 1.3 billion identifiers issued to date, saving an estimated \$27 billion.⁵⁹

The United States has made some headway in this direction. Login.gov was set up so that people can interact with multiple agencies using one login account and currently has over 100 million users across the country.⁶⁰ A general purpose form-builder that thousands of government services could use as a shortcut is in development.

That is much better than nothing, and giving that team support and funding will help. But the general sense is that we are behind, and left as is, the United States is likely to fall ever further back rather than catch up.

Platforms and data infrastructure provide a set of building blocks that service teams can use with very little additional effort. The more infrastructure that exists, the quicker and cheaper it becomes for teams to run more experiments with different policy choices, at scale. Teams no longer have to build everything from scratch.

CASE STUDY: DIGITAL RESTART FUND, AUSTRALIA

The New South Wales state government set up a Digital Restart Fund with an initial \$100 million budget in 2019 to build consistent, reusable platforms across government, improving digital services in the process. The fund is now \$2.1 billion, and funds iterative, multidisciplinary approaches to planning, designing, and developing digital products and services.

The fund has been intentionally established to support more incremental, product-type approaches to investment in public services and capability. Project bids are capped at AUS \$20 million, so that megacontracts for long periods of time are off the table, and the presumption of adapting delivery objectives to what is learned from testing new services in reality is built in. Project bids are prioritized according to centrally set criteria, including whether they align with particular outcomes or missions the government is focused on.

By the end of 2022/23, the fund had invested in nearly 300 transformation projects across NSW's portfolio of agencies. On average, the return on project investments was nearly three times the project cost.⁶¹

RIGHT-SIZE AI GUARDRAILS

We've discussed the "cascade of rigidity" that occurs when well-intentioned laws and regulations become increasingly inflexible as they step down through bureaucratic hierarchies. This dynamic disproportionately affects technology issues in part because few people in that hierarchy understand the technology in question, and lack of understanding drives towards a "better safe than sorry" approach. We see this, for example, in how the security requirements in the Federal Information Security Modernization Act, meant to be a flexible framework, have become an exercise in mandatory checklists that actually reduce security rather than enhance it.

59. Tom Loosemore, "What we mean when we say 'Be Bold,'" Public Digital, October 16 2024, <https://public.digital/pd-insights/blog/2024/10/what-we-mean-when-we-say-be-bold>.

60. "Frequently Asked Questions (FAQ)," Login.gov Partners, last accessed December 2024, <https://www.login.gov/partners/fag/#:-:text=How%20many%20people%20have%20signed,300%20million%20sign%20Dins%20annually>.

61. "Digital Restart Fund: Success Stories 2022/23," NSW Government, 2023, <https://www.digital.nsw.gov.au/funding/digital-restart-fund/2022-2023-success-stories>.

This dynamic is at risk of negatively affecting government's adoption of AI. The Biden Administration's AI executive orders and guidelines, while reasonable on paper, are likely to become overly restrictive as they are implemented. Requirements like public consultation and equity studies could effectively prevent AI adoption due to lengthy processes.

The solution is greater internal competency and capacity, not more rules. Agencies need more flexibility and discretion rather than rigid compliance requirements. This includes improving trust between oversight bodies and agencies and creating working capital funds and other mechanisms to give agencies more flexibility, as discussed elsewhere in this paper.

Why is most policy educated guesswork with a feedback loop measured in years?

- Tom Loosemore, Public Digital

4. Close the loop around outcomes

We need a new operating model for government if we are to restore our capacity to achieve our policy goals. This model must close the open loop we described in Part 3: a one-way system from law- and policy-making to implementation to real world outcomes that offers little space for learning and adjustment along the way. We can no longer rely on media coverage and elections, blunt tools that tend to be saved only for the most catastrophic errors, as the main corrective mechanisms.

Closing the loop means that we must apply test-and-learn approaches. This means conducting multiple small-scale experiments at the boundaries of policy and delivery — and doing this permanently, in pursuit of a policy intent or outcome. Incremental changes are scaled up once there is good evidence they work in reality. Test-and-learn does not mean simply running lots of pilots. A pilot implies starting with a phase for learning, which then ends as you move into “roll out.” Responsiveness is an embedded attribute, not a phase on a timeline. Closing the loop means the learning doesn't stop at an arbitrary moment.

CASE STUDY: UNIVERSAL CREDIT, UK

Universal Credit (UC) represented the biggest reform of the UK's welfare benefits system since 1948, combining six working age welfare payments into one and comprehensively overhauling the technological, operational, and policy underpinnings at the same time. It was the government's flagship domestic policy. Yet after three years, the program had spent £425 million, gone through five program leaders, and had not delivered a working service to a single claimant.

UC is a rare successful example of turning around a huge program that was “heading for nowhere but the rocks.”⁶² A total reset of the service design and delivery approach moved away from the usual linear, waterfall program to one that more intentionally took a test-and-learn approach to close the loop.

Several factors contributed to the ultimate success of UC. A brand new team was set up, initially just 15 people, composed of a wide variety of disciplines: policy, economics, service design, software development, user research, and frontline operations. It was led effectively by a senior trio, with one policy expert, one technologist, and one operational leader. This team was charged with something quite different from the marching orders given to implementation teams in the U.S. The minister in charge told them: *“I want you to deliver an intervention that means we support more people to find more work, more of the time, while protecting those who can't work.”* Note that this was not the same as: *“I want you to deliver a set of written directives called Universal Credit.”* In effect, the specifics of the policy would be worked out as they went.

UC was initially tested as a complete end-to-end service with a pilot group of just 100 claimants in a single, small postcode area in south London. This area was carefully selected to test whether the assumptions the team had made about the core proposition were correct. This early test was made possible by tabling secondary legislation (see above) to allow for legal experimentation within a tightly-defined part of the country. Doing so revealed unanticipated challenges within weeks, such as how payments information was displayed to claimants, or what the definition of a “couple” was (something that had meaningful policy and legal consequences). Over time, the team tested the core proposition with larger groups. Next, the team started to test assumptions about how that proposition could be scaled nationally. Rather than a linear process of writing policy at the start and living with the consequences, the team started with outcomes, drafted a design, and continually iterated based on what they learned.

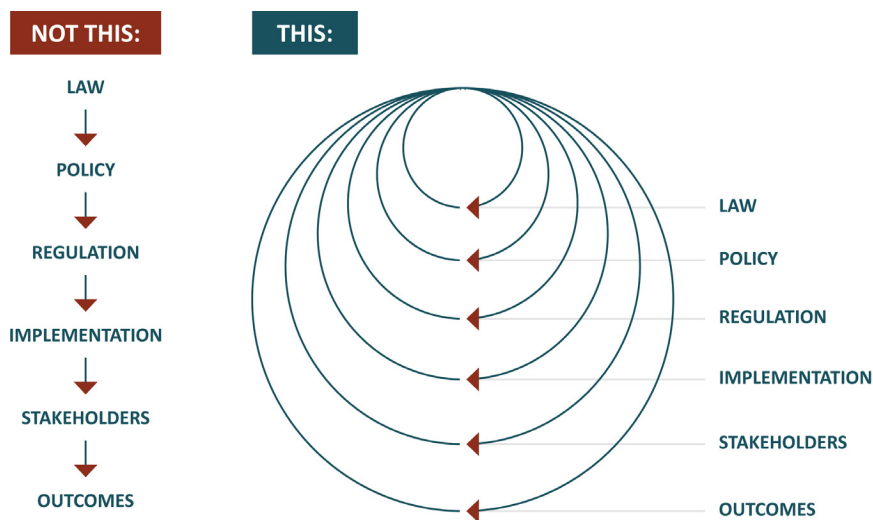
To picture more clearly how this worked, imagine the moment the UC team showed the minister responsible for the service a video of some user research they'd done with a couple enrolled in UC. The video made it clear that the policy assumptions about how Universal Credit should operate when both members of a couple are in the program were faulty. “Okay, I recognize that's wrong,” he said. **On the spot, the minister adjusted the policy accordingly.** Because they were still piloting the program with such a small number of people, no code had been determined and no operational training had happened, so it was easy to adjust. They learned early what most programs fail to learn until the costs to change are too high. But most importantly, because the policy-implementation loop was so tight, they could act on it.

Today, Universal Credit serves over five million households across the UK,⁶³ and has become a world-leading example of how to build successful user-centered public services at scale. It was a dog that didn't bark during the Covid-19 pandemic; despite a precipitous increase in uptake, the service responded when it was needed most. That was because it had been built through a way of working that allowed multiple changes to policy directives to be made in days, not years. This is what you get when you close the loop.

62. See here for a comprehensive and independent view on the program's early days and subsequent turnaround: Nicholas Timmins, “Universal Credit: From disaster to recovery?” Institute for Government, September 5, 2016, <https://www.instituteforgovernment.org.uk/publication/report/universal-credit-disaster-recovery>.

63. “Households on Universal Credit Dashboard,” UK Department of Work and Pensions, last updated May 2024, <https://stat-xplore.dwp.gov.uk/webapi/metadata/dashboards/uch/index.html>.

When we plead for “closing the loop around outcomes,” we actually have many loops in mind. Agencies must be able to test and learn with their own stakeholders within their own sphere of authority, closing the loop on the insights they gain to improve outcomes. Within agencies, this requires far greater integration of implementation teams (and often frontline workers like inspectors, adjudicators, case managers, and warfighters) into policy-making spheres, so that agency-level policy and regulatory activity benefits from the feedback gained by real-world experimentation. (See the Universal Credit case study as an example.) Where those learnings indicate it, they need to go back up the chain not just to executive branch policymakers and oversight bodies, but to lawmakers in Congress, to close the loop between the original intent and the outcomes resulting from the prescribed activities. In this regard, agencies can play a role in closing the feedback loop between elected officials and the constituents they represent.



To close the Congress-to-agency loop requires a transformation of oversight. The good news is that such a transformation is already under way, whether people like it or not, as a result of the Supreme Court’s June 2024 *Loper-Bright* decision. That ruling largely reversed what has been known as *Chevron* deference, the doctrine that instructed courts to defer to an administrative agency’s interpretation of ambiguous statutes as long as the interpretation was reasonable.

While the overall impact of the decision is yet to be seen, it is likely to greatly curtail the policy effectiveness of any future administration that seeks to act without clear authority from Congress.⁶⁴ That means that Congress itself will have to either step up and assert its authority or cede power to the courts.⁶⁵ And “stepping up” will require the legislative branch to invest in its own capacity⁶⁶ and establish more robust, continuous feedback loops with the executive branch.⁶⁷ Executive branch agencies will need to meet their legislative

64. Marcia Brown, “The Supreme Court’s recent decisions could undo big Biden accomplishments,” *Politico*, August 26, 2024, <https://www.politico.com/news/2024/08/26/chevron-biden-harris-legacy-00176268>.

65. Satya Thallam, “Congress in a Post-Chevron World,” Foundation for American Innovation, July 23, 2024, <https://www.thefai.org/posts/congress-in-a-post-chevron-world>.

66. Taylor J. Swift, “Congress must get serious about its capacity or cede power to courts,” *The Fulcrum*, August 21, 2024, <https://thefulcrum.us/governance-legislation/congressional-capacity>.

67. “Chevron Changes 101: What it Means for Your Job Now,” POPVOX Foundation, July 19, 2024, <https://www.popvox.org/events/chevron-changes-101>.

counterparts in this task: they will no longer be able to make their best (and often wrong) guesses about legislative intent and hope for the best, as those guesses are far less likely to stick. Neither branch may be asking for ways to close the loop, or at least not in that language, but in a post-*Chevron*, fast-moving, and resource constrained world, both very much need it.

The new models we must build will have many facets, not all of which we can touch on here. But we can start with ways to realign both branches around outcomes instead of outputs by changing the way funding and oversight are conducted.

ALIGN FUNDING TO SUPPORT FEEDBACK LOOPS

One of the many things that makes it hard for government to close the feedback loop is how it tends to spend. In Part 3, we talked about what happens when the public sector goes beyond buying the goods and services it needs and instead abdicates responsibility to its vendors. In a section above, we described a contrasting model for procurement particularly of digital work, one that gets far better value from the contractor ecosystem because it puts internal competency and capacity first and allows for continual improvement over time. These discussions all lead toward a critical, and increasingly obvious, conclusion: many of the dysfunctions that plague government are downstream of our funding practices.

At the heart of so much of our dysfunction is a shared lie. No one admits it's a lie, but everyone knows it is. For any reasonably complex (or even complicated) need government may have, we pretend we know what meeting that need will cost. Agencies and the White House Office of Management and Budget pretend when they submit budget requests, teams pretend when they do their annual planning, vendors pretend when they answer RFPs. Congress pretends to believe these requests...sort of. Members also like to point out that they don't trust them, based on the past history of cost overruns.

Obviously, there's a shared awareness that budgets are best guesses, and that actual spending may not match perfectly with the plan. But what makes the way we budget a lie is the terms of the conversation we have around it. In government culture, the best budget request is the most detailed, the most specific, the most certain of the future. It is granularly detailed because it is assumed to be informed by the most careful planning. The worst budget request is one with round numbers. A colleague was once dressed down for submitting a three-year budget that had only zeros after the decimal point. Rounding to the dollar meant that the budget hadn't been built pen-by-pen, computer-by-computer, salary-by-salary...for three or more years out.

In the real world, the best planning does not mean budgeting pen-by-pen. The best planning tests the riskiest assumptions first and is fastest to adjust based on those learnings. (Much catastrophic failure comes from leaving the hardest parts to the end, when too much has been locked in and there's no time to change. Note that contractors are highly incentivized to do exactly that, as they will get paid for all the easy work up to the end, and then charge for enormous change orders when the hard part ends in failure.) The best budgeting supports that testing and learning so the team can focus on getting the best possible outcome within the spending limits. It is extremely hard to test and learn if the funding for your project is contingent on pretending you know everything that is going to happen years into the future. Not only do we make public servants pretend to have crystal balls, but we punish them when their predictions turn out to be wrong.

The move to cloud computing provides a perfect illustration of how this facade of certainty hobbles our effectiveness. One of the key advantages of cloud computing is the flexibility it offers. Instead of having to

guess how much infrastructure you'll need well ahead of launching a website and buying what you hope is the right number of servers and the sufficient bandwidth, you can essentially rent a flexible amount of capacity from a cloud computing provider and only pay for what you use. If traffic is less than expected, you save money. If it's more, you pay more, but at least your website stays up as the cloud provider seamlessly handles the extra load. And you don't have to buy a lot more than you will generally need just because loads will occasionally spike (which is extremely common with government websites that tend to get used only at, say, tax filing time, or when a national emergency strikes).

An engineer at the General Services Administration tells the story of working with a contracting officer to try to purchase cloud services about 10 years ago. The engineer had prepared enormous amounts of data in advance of meeting with the contracting officer, but in the meeting, he found that he lacked the one piece of information the officer needed: how much these cloud services would be used. The officer could not put out a bid to procure a service if he didn't know how much he would be asking for. The engineer couldn't tell the contracting officer how much "cloud" he needed to buy, because not knowing is exactly the point of this technology. The two went around and around, respectful of each other but increasingly frustrated. Finally the contracting officer took a deep breath and said, "Let me explain how contracting works in the U.S. government. We put in an order for 100 sandbags, we get 100 sandbags." And the conversation was over.

Government has been trying to move to the cloud since Federal CIO Vivek Kundra's 25 Point Plan to Reform Federal IT in 2010.⁶⁸ It's no wonder the move has been slow, given the number of conversations that must have ended up with sandbags. But it's also true that the cost savings that were expected from this move have largely failed to materialize. The private sector saw significant cost savings by moving to the cloud, but NextGov recently reported that the federal government could be paying \$500 million a year more than it needs to for cloud services. This is not due to the problems the GSA engineer encountered, which have been mitigated over time, but rather the fact that "cloud service providers typically offer significant discounts for contractual commitments of three years or more...The Anti-Deficiency Act mandates that the government cannot pre-commit to out-year expenditures prior to Congress allocating funds. In most cases, with some exceptions, this means that government buyers of cloud can only commit annually to procuring cloud services, one year at a time."⁶⁹ Add to this constraint all the custom requirements government imposes far above and beyond what private sector customers ask for, and government pays so much more for cloud services that in many cases moving to the cloud has cost more, not less.

Savings of that scale could be helpful in what may be an upcoming era of fiscal austerity. But that is a tiny portion of the value to be gained from a hard look at the funding practices of government. Funding that enables test-and-learn approaches and provides teams sufficient flexibility also allows authorizers, appropriators, and oversight bodies to hold agencies accountable to meaningful outcomes instead of largely fictional plans. Taxpayer dollars are always at risk of waste, fraud, and abuse, but the rigid, falsely certain, and never-guaranteed funding streams we offer today do little to protect against these harms. They do a lot to keep us from getting the value for the dollar that taxpayers deserve.

68. Vivek Kundra, "25 Point Implementation Plan to Reform Federal Information Technology Management," U.S. Chief Information Officer, December 9, 2010, https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/egov_docs/25-point-implementation-plan-to-reform-federal-it.pdf.

69. Michael Garland, "Is the Government Losing \$500M a Year in Cloud Savings?" NextGov, June 6, 2024, <https://www.nextgov.com/ideas/2024/06/government-losing-500m-year-cloud-savings/397152>.

FUND INTERNAL COMPETENCE FIRST

Ongoing, flexible funding is only part of the picture. We also need to fund the right internal competencies and capabilities. Nothing in the product model — not the early inquiry, nor the scaling up, nor the hiring of vendors — works without staff with the right skills and approaches. What competencies and capacities must government develop and retain in-house, and what is most appropriate to outsource?

The question isn't a binary decision over whether products and services should be given to vendors or developed in-house. The question is whether the agencies responsible for them have the appropriate internal capacities to understand their own needs, craft a meaningful vision for the product or service they need and how it will integrate with their agency operations, and hire and manage both internal and external resources to deliver on that vision in ways that allow for ongoing learning and adjustment. While many agencies are making laudable progress, in far too much of government the answer to that question is a decided no. Absent that competency, we hand over to highly paid consultants and contractors what should be inherently governmental work.

Again, software provides a useful example. Proper leadership can often be established with a relatively small internal team: researchers who can conduct discovery sprints to better understand needs, a product manager who can develop a vision for how the product will work and integrate with agency functions, lawyers and domain experts who can translate between the policy environment and the software design, procurement officers with expertise in developing RFPs for agile, iterative development services that can select among bidders on the basis of demonstrated competence, and a high-level product owner with sufficient authority within the agency to make decisions, even when they involve trade-offs. A modestly sized, cross-disciplinary team like this can take what would otherwise have been a “one throat to choke” megaproject with a price tag in the hundreds of millions or even billions, and break it down into manageable parts that deliver results better, faster, and cheaper.

ALIGN OVERSIGHT TO NEW MODELS

We need to shift to incremental, ongoing funding for products and capabilities supported by the right internal competencies. But changing the funding models will fail if the bodies responsible for oversight of these funds expect the same artifacts and trappings of the legacy models. You can't oversee agile, user-centered development with requirements-driven waterfall frameworks. Everyone just talks past each other.

Elected leaders often bemoan the risk aversion of the bureaucracy, wishing that the public sector workforce would more consistently take common-sense, expeditious steps towards defined goals rather than rely on the precedent of established procedures when the two are in conflict. You hear this complaint frequently in Congress as a befuddlement about why things that have been ordered in law take so long to come to fruition, like why the energy infrastructure hoped for in the Inflation Reduction Act is still stuck in permitting processes, or why so little of the investment authorized in the CHIPS Act has been awarded. The remedy should be oversight, which is underused and much needed to keep implementation on track. But the way it is conducted today provides little opportunity for mid-course correction and tends to increase risk aversion, not reduce it. Oversight is in desperate need of an overhaul.

First, let's address how oversight tends to increase risk aversion in the bureaucracy. Members of Congress and legislative staffers try to alleviate the burdens of established procedures; for example, they will grant

special authorities to agencies, like streamlined hiring or procurement pathways. But they then find that these authorities are rarely used. The reason is that though it may be legal to use these authorities, their legality may not protect those who employ them if it is later determined by an agency compliance officer or inspector general that there wasn't sufficient justification to deviate from what is considered standard operating procedure, and it is not the legislators who will make that call. Public servants remain risk averse in part because they fear internal oversight that will hold them accountable to the wrong things.

Rigid adherence to protocol is critical in many circumstances, and can be lifesaving. But in an increasingly complex, fast-changing world, it can also be very much the wrong stance. The Navy distinguishes between the modes of "front of sub" and "back of sub." The back of a nuclear submarine, where the nuclear missiles live, is run by the book. You don't deviate from the checklist. You don't innovate. You don't question. The front of the sub, on the other hand, is responsible for navigating through dark waters. You have to improvise. You have to make judgment calls. There are manuals and checklists, for sure, but the nature of the work calls for a different approach, and the Navy recognizes that the cultures of front and back have evolved appropriately to meet distinct needs. Too often, we conduct oversight as if all of government was "back of sub." Oversight can also be weaponized by applying "back of sub" standards to "front of sub" circumstances. Public servants are painfully aware of this possibility.

Excessive risk aversion drives much of the core dysfunction of government. Oversight, as often practiced, increases risk aversion in at least three ways. First, oversight focuses almost exclusively on what goes wrong. This gives public servants a very clear picture of all the ways they could fail (or get in trouble), but little understanding of what it looks like to succeed. Second, when oversight bodies examine what's gone wrong, they closely examine any deviation from standard process and procedure, even if that deviation is legal. Fidelity to established process is often at odds with achieving the intended outcome, but public servants choose process over outcomes in large part because oversight will punish them for choosing otherwise. When public servants say, "There is no upside, only downside," this is in part what they mean. Third, the previous two dynamics encourage ever greater proceduralism, the establishment of ever more processes to validate the decisions of public servants in the hopes of insulating them from blame. A more outcomes-oriented government cannot be achieved without significant changes in the approaches and practices of its oversight bodies.

Oversight in effect comes from three primary sources. Executive branch oversight includes the 74 Inspectors General within federal agencies. In the legislative branch, Congress conducts hearings and relies on the Government Accountability Office to set standards, conduct audits and investigations, and produce reports. And, though not part of government, the media serve an oversight function, and their output affects all three branches of government. We will address only the first two here.

Across both branches of government, there is little that's needed in terms of statutory change. Oversight could become a driver of greater effectiveness, with more upside and less downside, primarily through changes in *practice* by those who engage in it. In the executive branch, that starts with convening Inspectors General to educate them about the downstream effects of their actions. Where "front of sub" circumstances exist, IGs must apply "front of sub" standards, appropriately accounting for the use of judgment. Today, many IGs feel compelled to list deviations from standard operating procedure even when they believe the deviation to be a perfectly reasonable "front of sub" choice. This inevitably results in an end to those deviations, which might otherwise have been seen as much-needed innovations, even if the IG had no intention of killing them. IGs need to do their job to hold agencies accountable, but they need to learn to do that job in ways that don't make the use of discretion by public servants so hard. It is that discretion that is necessary to get the better outcomes IGs seek.

Congressional oversight is subject to the same conundrum, but the dynamics are a little different. Members of Congress and their staff are unlikely to notice minor procedural violations, engaging mostly when something has gone substantively wrong. And they occasionally use the opportunity of failure to address some of the underlying causes, as when they try to provide agencies greater flexibility when rigid rules are holding them back. But they only engage when something goes wrong, or is perceived to, contributing to the reality of “all downside, no upside” for public servants. Imagine a company in which management only spoke to employees for punitive purposes. Not only would employees become enormously defensive, they would also have no model for success other than following extremely prescribed rules. Similarly, though Hill staffers and members privately appreciate public servants who take appropriate risks, there is no public expression of that praise. If Congress were to flex a portion of its oversight muscle finding and lifting up positive examples of public servants putting outcomes over process, it would go a long way to reorienting bureaucratic behavior.

One reason oversight only seems to happen when something goes wrong is that few people realize it is going wrong, or how seriously wrong it is going, until late in the game. It’s not just that Congress doesn’t know it’s not working; it’s often that more senior officials in agencies don’t know themselves. Consider President Obama’s statements that healthcare.gov would work in the days before its disastrous launch. The team responsible had demoed it to him personally, and it seemed to work — as long as there were no other users on the site. To the best of his knowledge, the site was going to function. But what appeared to be a misrepresentation infuriated lawmakers who felt deceived.

At least in this case there was a demo. Very seldom does oversight involve reviewing anything other than paper— spending that has gone over budget, timelines that have stretched as key milestones have been missed. These metrics essentially measure how well a planner predicted what would happen before the team started working. They don’t reflect what leaders really need to know: are the riskiest assumptions being tested, and if so, what is the team learning? How is it adapting to what it learns so it can achieve the intended outcome? How are the people the project is supposed to serve involved, and how are they reacting? Checking requirements or steps off a list that was defined long before the project got going tells us little about whether the project will succeed, and yet that is where oversight discussions tend to focus. Instead of “show us how you are performing to plan” the questions need to be “show us what you are learning and how you will need to adjust to succeed.”

A new initiative at the Department of Veterans Affairs (VA) emphasizes practical skills over written proposals when selecting vendors for digital services contracts.⁷⁰ This project, known as SPRUCE, uses “tech demos” where vendors are asked to demonstrate their abilities through hands-on tasks rather than simply writing about them. This “do, don’t say” approach reflects an important shift in procurement practices towards valuing actual capabilities. Oversight bodies should learn from this practice, asking for demonstrations of genuine progress over written reports.

FOCUS ON ENABLEMENT OVER MANDATES AND CONSTRAINTS

Recently one of us was asked by Congressional staffers to advise on how to improve the performance of an agency they believed was failing to carry out Congress’s will. After an overview of initial ideas, we offered an observation. “It sounds like you’re searching for the right mandates and constraints to impose on this

70. Steve Kelman, “VA acquisitions asks vendors to show their work,” *NextGov*, October 24, 2024, <https://www.nextgov.com/ideas/2024/10/va-acquisitions-asks-vendors-show-their-work/400530/>.

agency to get the results you want.” The staffers agreed that was a reasonable characterization of the request. “But this agency has been on the receiving end of mandates and constraints from this body for decades,” we reflected. “They don’t seem to be working.” Again, the staffers agreed. In fact, they shared, the agency’s performance only got worse with every attempt to make it better.

When an approach fails so consistently, it is time to reconsider the approach itself. Seen from the perspective of the agency, each mandate from on high contributed to an undifferentiated and ever-growing clutter of work to do, difficult to prioritize. Each constraint imposed, while seemingly reasonable in its own right, added up over the decades to a sense of paralysis. One agency leader described her relationship to the Hill as a game of Mother May I, in which she has to ask for permission before taking any action. It’s like trying to run a marathon, she explained, but with each step forward requiring permission from Congressional committees. Move before the permission is granted, and you’re sent back to the start.

Crafting mandates and constraints is a key way that legislative bodies ensure their intent is followed. But it is not the only way. Legislators are also charged with oversight of the laws they pass, a muscle they tend to flex only when something has gone very wrong, and could flex very differently. And legislation does not need to take the form of adding mandates and constraints. It can remove them. It can send agencies the message that one thing is a priority by removing another thing. It can provide flexibility like working capital funds, to allow for experimentation before scaling up. It can reduce unnecessary reporting requirements, or harmonize those requirements with those of other bodies, so that agencies spend less time reporting and more time doing.

It is Congress’ job to define *what* the goal is and provide leeway for agencies to use the right *how* to accomplish it. Legislators can do more to ask the agencies what might help get the results that Congress wants and to build in affordances for ongoing dialogue and adjustment – closing the feedback loop.

This may sound enormously naive. Today, the trust required for either Congress or the executive agencies to engage in this kind of dialogue seems light-years away. If Congress (or the White House for that matter) doesn’t want the agency to succeed in its mission, then the above scenario is irrelevant. But flying under the radar of all the political disagreements we could find ourselves caught in, there will always be people in government trying to figure out issues of performance and effectiveness, trying to authorize spending with some degree of confidence that the spending will show results. Government simply does too many things for all of them to be controversial. Better oversight could begin to flourish out of the spotlight.

Even absent partisan differences, honest problem-solving dialogue between the branches is hard not only because Congress defaults to adding mandates and constraints, but also because agencies clam up and become enormously defensive in their relationships with the Hill. Anyone who’s ever been on the receiving end of Questions for the Record in an agency will tell you the direction their legislative affairs team gives them is “answer quickly and say nothing.” Even if the people responsible would like to engage, agency gatekeepers fear the weaponization of any substantive information, and allow as little as possible through. Behavior change would require bilateral disarmament, so to speak. While that sounds far-fetched, the alternative (the status quo) serves neither branch. At some point, even far-fetched alternatives become worth trying.

“Hierarchical, one-directional, top-down communication is built into the structure of the entire country...But even the people issuing the orders are fed up with the outcomes of this dysfunctional waterfall. It serves no one.”

Recoding America⁷¹

INVEST IN CONGRESSIONAL CAPACITY

With roughly 1/120th the staff of the executive branch,⁷² Congress lacks the capacity and expertise to address every detail of implementation detail. In a post-*Chevron* world, Congress must be very clear about exactly what delegated authority it assigns to agencies. The Hill will need to learn to engage with executive branch agencies as partners by establishing a well-defined dialogue between Members of Congress and agency personnel. This will require closer oversight and understanding of implementation challenges to identify opportunities to refine policies as issues arise.

Congress will need to strengthen its own capacities. It must bolster its workforce, especially for committees and institutional offices like the Office of the Clerk, the Congressional Research Service, the Government Accountability Office, and the Office of Legislative Counsel. But it must also upskill existing and new staff members, not only to understand how to write law that will not invite legal challenge under *Loper-Bright*, but to write law that is fundamentally implementable by agencies. Congress is used to being at the top of a waterfall. That will no longer work. Now it needs to be in a dialogue with agencies in a consistent feedback loop.

Congressional experts have proposed a variety of approaches for improving the legislative branch’s ability to monitor and respond to regulatory developments:⁷³ from beefing up committee staff with technical and legal experts on the programs within their jurisdiction⁷⁴ to establishing a dedicated Congressional Regulation Office (CRO).⁷⁵ In addition to human expertise, the legislative branch should give itself modern tools and access to data for ongoing monitoring and oversight. In recommendations for changes to House Rules in the 119th Congress, POPVOX Foundation proposed the creation of a chief data officer for the House that would “prepare a strategic plan to provide committees with continuously updated data, enabling them to better monitor programs within their jurisdiction and respond to emerging issues in real time.”⁷⁶

71. Jennifer Pahlka, *Recoding America* (New York: Metropolitan Books, 2023).

72. Taylor J Swift, “Congress must get serious about its capacity or cede power to courts,” POPVOX Foundation, August 21, 2024, <https://www.popvox.org/blog/the-fulcrum-op-ed>.

73. David Dayen, “How Congress Gets Its Groove Back,” *The American Prospect*, September 30, 2024, <https://prospect.org/politics/2024-09-30-how-congress-gets-its-groove-back/>.

74. Jim Saksa, “More congressional staff? Conservatives are coming around to the idea,” *Roll Call*, July 30, 2024, <https://rollcall.com/2024/07/30/more-congressional-staff-chevron-deference/>.

75. Philip Wallach and Kevin R. Kosar, “The Case for a Congressional Regulation Office,” *National Affairs*, Fall 2016, <https://www.nationalaffairs.com/publications/detail/the-case-for-a-congressional-regulation-office>.

76. “POPVOX Foundation Encourages the House to Create a Chief Data Officer, A Joint Committee on Continuity, and Simpler House Schedule When Adopting Rules for the 119th Congress,” POPVOX Foundation, August 20, 2024, <https://www.popvox.org/blog/119th-house-rules-recommendations>.

INCENTIVIZE LEGISLATORS TO CLOSE THE LOOP

Arguably, the biggest impact legislators can make for their constituents is helping the systems that administer our current laws, policies, and programs run better. Any constituent services staffer on the Hill will tell you how much frustration their callers express with their disability claims, their tax returns, their Medicare coverage. Congress passes bills every year that affect how these programs function, but as soon as they're passed, the focus moves to what should be legislated next. The bulk of a legislator's time is spent not on improving the delivery of the programs we have today, but on writing new rules to add to the old ones, and introducing entirely new topics.

Part of the reason for this is that introducing bills gets them attention. “How a bill becomes a law” has a clear narrative and defined cadence of activities. Will a bill be introduced? Get out of committee? Make it to the floor? The press and advocacy organizations, including their members and donors, orient their own workflows and outputs around these events, creating a major avenue through which voters learn about the legislator's work. And legislators need voters to know about their work. Oversight should offer opportunities for attention as well, but, in practice, it largely occurs in response to high-profile failures when what's expected — but rarely helpful — is outrage from Congress.

But reformers recognize the need for these same legislators to spend far more time following up on existing legislation that's not all outrage — the normal course of checking on implementation, understanding where it may be stuck, what unintended consequences might be emerging, and observing whether it's having the intended impact. These are the feedback loops we've discussed earlier.

To make that happen, some are starting to experiment with workflows that create narrative vectors for oversight activities. One state is hoping to pilot a legislative tool that helps members and stakeholders evaluate and improve the impact of enacted legislation. The tool features a consistent cadence of follow-up and oversight activities over a calendar year, including announcements of intended targets, roundtables and other engagements with agencies, constituents, academics, and advocates. Each bill or issue targeted would ultimately conclude with either suggestions for fix-it bills, helpful executive actions, or a summary of the positive impacts the bill has had so far. By mimicking the cadence of activities that the members, advocates, and press already understand, and that attract their attention, leadership hopes to redirect some members' activities away from new initiatives and towards follow up. Congressional leadership could follow suit.

Another promising example is the embrace of an innovative digital system by the Chief Administrative Officer of the House of Representatives. Under development by the House Digital Service, the Casework Data Navigator will draw insights from aggregated constituent data to inform improved oversight.⁷⁷ Efforts like this open new opportunities to identify problems for proactive solutions.

MAKE NEW OPERATING MODELS THE NORM

Taken together, the steps outlined above are more than a checklist of tasks to be completed. They represent the start of a different operating model for government, one that may feel new but is in fact far more consistent with the vision our Founders held. That vision has been blurred as the nation grew, the government took on more and more responsibilities, and the ever-growing complexity of our bureaucracy created vast

77. “POPVOX Foundation Shares a Roadmap for Congressional Casework Data,” POPVOX Foundation, July 7, 2024, <https://www.popvox.org/blog/roadmap-for-casework-data>.

distances between the policies as envisioned and the lived experience of those policies in the wild. The new operating model must leverage the tools and thinking of the digital age to reverse this unwelcome trend and collapse distances in a number of directions, particularly between policymakers, implementers, and the public whom both are meant to serve.

A shift to a new operating model that can connect across branches of government will be the work of more than one Congress, more than one presidential administration. It will require the leadership of more than one party. It should begin now.

Conclusion: What a different *how* offers

When a section of I-95 collapsed in 2023, Pennsylvania Governor Josh Shapiro knew that how we fix infrastructure today simply wouldn't cut it. He invoked emergency powers and directed his team to use them. Tasks in procurement, contracting, negotiations, clearances, staffing, and the like that under normal circumstances could have taken years were completed in days. The interstate highway was open again in under two weeks.

When President Trump announced Operation Warp Speed in May of 2020, few thought we would have vaccines against the Covid virus within a few years. Vaccinations in the United States began on December 14 of the same year.

When President Biden announced at the start of his term that free Covid tests would be mailed to every American who wanted them, a website to take those requests needed to be built. Websites traditionally take years to stand up in government, given the many months needed for sequential stages of procurement, development, and numerous compliance clearances. Covidtests.gov went live, serving millions of users, in less than a month.

In each of these cases, the quality of work does not seem to have suffered from the speed with which it was done. In the case of the website, and possibly the interstate, cost goes down with speed, not up; speed saves countless staff and vendor hours, even if the concrete or some of the labor comes at a premium because it is needed faster. The old adage that says you must pick two from fast, cheap, and good gets complicated by the dysfunctions of bureaucracy in which going slower gives you outcomes that are worse *and* more expensive. These outcomes raise important questions about the tradeoffs we make when we slow processes down in the pursuit of quality, safety, cost, and ethics. How much more of those are we really getting by going so slowly? What do heavyweight processes that seek to safeguard against harms really buy us, and what do they cost us? How might we more thoughtfully rebalance these factors?

We are already operating in a new *how*, albeit dramatically unevenly. The examples above are all born of crisis, and crisis is indeed the biggest driver of new ways of leading and working. Other instances don't leverage the same pressure of daily headlines, but simply recognize the urgency of the public need. There were no daily headlines shaming the State Department for its antiquated passport renewal process. And yet, the need to do this task in-person at a government office was not only inconvenient for millions of Americans and expensive to administer (and a major percentage of the Congressional casework load), but yet another small sign of government's failure to keep up with the pace of change. As discussed in the case study in Part 3, a U.S Digital Service team brought new ways of working to the Bureau of Consular Affairs, and the result is greater convenience for the American public and greater efficiency in the bureaucracy. More importantly,

Consular Affairs now has the opportunity to align their planning, funding, staffing, contracting, and reporting around a fundamentally different model of delivering value. They will need support, aircover, and strong leadership from both the executive and legislative branches in order to continue this transformation and see further benefits, some of which can be dramatic. That support and leadership is needed across government.

Part 5: Recommendations

HIRE THE RIGHT PEOPLE AND FIRE THE WRONG ONES

1. ACTIONS FOR THE WHITE HOUSE:

- a. Appoint a bold reformer to lead OPM
- b. Increase OPM delegation of authority to agencies for personnel actions
- c. Initiate wide-ranging workforce planning to rebalance “go energy” implementation roles and “stop energy” compliance roles
- d. Ensure each agency has high-level leaders with operational experience to drive policy implementation
- e. Revise civil service regulations, operations, and processes to truly honor merit system principles — prioritize hiring and accountability rules and practices
 - o Sunset the use of self-assessments for most positions within two years
 - o Revise the Delegated Examining Operations Handbook to ensure that hiring managers have appropriate involvement in the candidate selection process
 - o Reduce the number of pathways for protesting disciplinary actions against employees

2. ACTIONS FOR CONGRESS:

- a. Propose comprehensive civil service reform to better support merit system principles and maintain the independence of the civil service, updating the Civil Service Act of 1978

3. ACTIONS FOR CIVIL SOCIETY:

- a. Support civil service reform by exploring pragmatic compromises between interest groups that may stand in the way of meaningful reform

REDUCE PROCEDURAL BLOAT

4. ACTIONS FOR THE WHITE HOUSE:

- a. Adopt a cross-agency priority goal of reducing burdens on public servants, and assign it to the OMB Deputy Director for Management
- b. Direct agencies to propose policy, process, and regulatory clean up and bring requests to Congress when needed

- c. Simplify and clarify guidance where possible
- d. Support legislative Paperwork Reduction Act reform
- e. Ensure that appropriate technology, including LLMs, are available to agencies to support regulatory clean up

5. ACTIONS FOR CONGRESS:

- a. Initiate cleanup of outdated, vestigial, conflicting, and de minimis statutes, working collaboratively with agencies and civil society
- b. Pass PRA reform that reduces the burden on agencies and incentivizes best practices for burden reduction on the public
- c. Update the 1946 Administrative Procedures Act (APA) to reduce the surface area for frivolous legal attack

6. ACTIONS FOR CIVIL SOCIETY:

- a. Host events and briefings for Congressional and administration leaders on successful decluttering efforts at the federal, state, and local levels

INVEST IN DIGITAL INFRASTRUCTURE

7. ACTIONS FOR THE WHITE HOUSE:

- a. Submit budget requests to Congress that prioritize critical needs and opportunities for improving digital experts and common platforms
- b. Revise the laws, regulations, processes, and practices that impede agencies from operating in the product model
- c. Revisit the concept of “inherently governmental” for the digital age and issue guidance to agencies about needed internal capabilities

8. ACTIONS FOR CONGRESS:

- a. Increase funding for expert digital teams, such as U.S. Digital Service, Technology Transformation Service, and GSA’s 10x
- b. Invest in the Technology Modernization Fund to drive adoption of product model development
- c. Require agencies to demonstrate proper internal competencies to accompany budget requests
- d. Work with OMB and agencies to change the laws, regulations, processes, and practices that impede agencies from operating in the product model

9. ACTIONS FOR CIVIL SOCIETY:

- a. Advocate for agencies and offices improving digital expertise and using best practices for development and acquisition

CLOSE THE LOOP AROUND OUTCOMES

10. ACTIONS FOR THE WHITE HOUSE:

- a. Direct agencies to appropriately staff legislative affairs functions to reduce risk aversion, increase bandwidth, and allow for greater candor where needed
- b. Convene Inspectors General (IGs) to educate them about the downstream effects of their actions and use their discretion when documenting reasonable deviations from standard operating procedures

11. ACTIONS FOR CONGRESS:

- a. Upskill existing and new staff members
- b. Invest in Congressional capacity, particularly staffing committees and institutional offices with needed expertise such as the Office of the Clerk, the Congressional Research Service, the Government Accountability Office, and the Office of Legislative Counsel
- c. Initiate a working group to explore changes in oversight practices that could increase the ongoing communication and problem solving between the branches that will be needed to effectively govern in the Loper-Bright era
- d. Identify and lift up positive examples of public servants putting outcomes over process to reorient bureaucratic behavior
- e. Shift oversight from “show us how you are performing to plan”. The questions need to be “show us what you are learning and how you will need to adjust to succeed”
- f. Revise oversight practices to expect and request product model outputs over legacy signs of progress
- g. Use legislation to remove mandates and constraints such as fixed funding that prevents experimentation and unnecessary reporting requirements
- h. Explore creating a new legislative branch agency tasked with aiding Congress’s authority over implementation analysis and liaising between the legislative and executive branches to address ambiguities and ensure that interpretation aligns with Congressional intent
- i. Develop tools to help members and stakeholders evaluate and improve the impact of enacted legislation

12. ACTIONS FOR CIVIL SOCIETY:

- a. Study and promote new operating models for interbranch cooperation
- b. Hold government accountable to new models, not old ones
- c. Highlight and reward addition through subtraction in the press and think tank reports

ABOUT THE AUTHORS

Named the “[The Most Interesting Think Tank in American Politics](#)” by TIME Magazine in 2023, the Niskanen Center promotes policies that advance prosperity, opportunity, and human flourishing, guided by the belief that a free market and an effective government are mutually dependent.

We are forging a [new vision of American governance](#) that draws on ideas from across the ideological spectrum, an approach made possible by our independence from rigid political coalitions. Our ideological openness gives us credibility to work with a broad range of actors, meeting them on their own terms and helping them converge on policies they might support for different reasons. We thus aspire to be a voice for the hopes and ambitions of Americans who feel shut out by the hostility of our politics.

Public Digital is a global digital transformation consultancy, headquartered in London, that radically changes how organizations work so that they deliver excellent services and greater impact, even when the future is uncertain. It has advised more than 40 governments around the world, including the federal and state governments in the US. It works with international funders and multinational businesses, including the World Bank, UN Development Programme, Gates Foundation and Bloomberg Philanthropies.

Jennifer Pahlka is a senior fellow at the **Niskanen Center** and the Federation of American Scientists, and a senior advisor to the Abundance Network. She served as deputy chief technology officer of the United States under President Obama and a member of the Defense Innovation Board under both Obama and Trump. She founded the award-winning nonprofit Code for America, which she led for ten years. Pahlka is the winner of a Skoll Award for Social Entrepreneurship and was selected by Wired magazine as one of the people who have most shaped technology and society in the past twenty-five years. Ezra Klein called her book *Recoding America* “one of the best policy books I have ever read” and “the book I wish all policymakers would read.”

Andrew Greenway is a co-founder of **Public Digital**, and a former senior UK public servant. While in government, he was one of the founders of the Government Digital Service, leading work to create new digital guidelines and standards; a model of institutional change copied by many governments worldwide, including the US. He wrote *Digital Transformation at Scale*, a book commended by the former Prime Ministers of the UK and Australia, which has become a reference text for driving institutional reform around the world. He also co-wrote the influential Radical How report on mission-driven government, published in 2024.

Thanks to Ben Bain, David Dagan, Nicholas Bagley, Marci Harris, Geoff Kabaservice, Soren Dayton, Kodiak Hill-Davis, Robert Gordon, Jennifer Erickson, Charlie Anderson, and Kelly Taylor for their help in writing and editing this paper.