INTEGRATED BENEFITS

OUR VISION FOR AMERICA.GOV

AMIRA BOLAND & MARINA NITZE

SEARCHLIGHT INSTITUTE

CREATING A ONE-STOP SHOP OF THE FUTURE

IN BRIEF

Since 1993, every federal administration has promised Americans a one-stop shop for dealing with their government, and every one of them built a website instead. The websites shipped, mostly on time, and some of them were quite good. Meanwhile, the rules underneath — the ones that would actually make “a single interaction” possible — only got more complicated.

"Integrated benefits" and "one-stop shop" and “portals” have long described the same hope: weaving the safety net together so people can easily access all the help for which they're eligible, without having to fight five agencies for it, and with more efficient machinery behind the scenes. But these terms can refer to at least six very different ways of trying to achieve that goal.

We think of them as rungs on a ladder. The bottom rung is a basic link directory; you then move up to a screener/eligibility quiz, then a single site where you can see personalized information from multiple places, and then up to the ability to enroll in multiple programs at once with one form. These are the “lipstick on a pig” solutions. They are interface-only work that require no substantive technical, system, process, or even human training changes, and rely on the existing underlying eligibility, adjudication, and enrollment machinery. This is also where a majority of time and resources have been invested to date.

Those websites were never the hard part. What trips up most people is everything *after* applying: a request for a pay stub you already uploaded, a letter from your landlord, an interview you couldn't get off work for, or a recertification notice that went to your old address. A better form upfront doesn't touch any of that. But the next two rungs at the top of the ladder — a shared eligibility engine that can automate enrollment across programs, followed by proactive automatic enrollment — *do*, because they meaningfully fix the underlying processes instead of tidying the front end. These are the two rungs that can meaningfully change a person's experience for the better.

If the theory is that getting help to an unemployed worker quickly means they can recover quickly (and therefore have no more need for benefits), the top rungs are the level at which you would invest. Unfortunately, across over 30 years of effort, nearly all of the attention has gone to the bottom of the ladder.

These lower-rung efforts didn't fail to move the needle because people were bad at their jobs or didn't try hard enough. They failed because no one on the interface teams was able to change the outcome; because no agency (or level of government) readily shares or delegates authority over its own eligibility rules; because programs won't or can't share data or align definitions; because one federal program can have 53 (or more) different state versions; and because the laws of management cybernetics tell us a mere portal can't absorb the level of complexity that today's benefit lines contain. That complexity has to be absorbed somewhere else — usually landing squarely on the applicant.

The work that actually produces the goals a "one-stop shop" vision proposes — the ability to share your information once and receive everything to which you're entitled — isn't building a website. It's four things:

1. **Documenting** every program's eligibility criteria in a common format, at every layer from federal statute down to state and county policy;
2. **Establishing** and **enforcing** one authoritative source for every fact the government needs to make a determination;
3. **Reconciling** criteria / definitions so they can be answered by the authoritative source; and
4. **Validating** that the systems administering these programs actually implement the law.

Do those four things and we can reach the top rungs. More to the point, most determinations stop needing an application at all. If every criterion a program uses is one an authoritative source can answer, the program can answer it for everyone, on its own, the way the Social Security Administration enrolls everyone in Medicare at 65. Criteria that can't be answered that way should get moved up to a population level (e.g., free lunch for the whole school), rewritten to accept self-attestation, dropped, or (a much smaller set) handed explicitly to a human with the authority to adjudicate.

With published criteria and a test suite every system has to pass on the back end, now anyone can build an accurate front-end interface, including the AI assistant to which a growing number of people are already delegating their daily tasks. People are *already* asking their AI whether they qualify for SNAP. The choice isn't whether AI sits between people and their benefits; it's whether the rules AI leverages are authoritative and tested.

Get the rules right and we could have a million one-stop shops, customized for each person's needs, and the government never has to build, procure, or maintain a single one of them. Instead, government can focus on the things only it can do: say what the rules are, verify facts, and pay.

None of this happens if we ask nicely. It requires a commanding presidential directive to hold agency heads accountable to this goal, to find every possible pathway within their power to get it done, and to clearly articulate necessary statutory changes, collaborating with Congress along the way.

This alignment and simplification of the safety net stands to be a uniquely transpartisan effort, because the status quo serves no camp's priorities. For conservatives who are more concerned about waste and fraud, adjudicating the same facts independently in every program is maddeningly expensive for federal, state, and local budgets alike. Scammers thrive in this chaos. Done well, this will combat fraud by making honest mismatches disappear, so the real fraud stands out. It also eliminates the overhead of running the same verification in a dozen places, and frees caseworkers from low-value processing to do the critical human judgment work only they can do. And for liberals who care about building a robust welfare state, this will get people what they need, fast enough to actually help.

Table of Contents

Executive Summary 5
The Six Rungs of "Integrated Benefits" 11
30 Years at the Bottom of the Ladder: A History of One-Stop Shops 18
Eight Lessons From 30 Years of Integrated Benefits Delivery Failures 23
- 1. Nobody Gives Up Power 24
- 2. Programs Will Not Share Data or Trust Shared Data 25
- 3. A Shared Eligibility Engine Is Really Two Problems 27
- 4. There Are Really 53 SNAP Programs 28
- 5. The Code Does Not Always Match the Law 29
- 6. Ashby's Law of Requisite Variety 30
- 7. A Front Door Is Still a Pull, Not a Push 31
- 8. Combined, Dynamic Applications Alone Can’t Work 32

Lessons From What’s Worked 35
How to Climb the Ladder 38
- Step 1: Consistently Document Eligibility Criteria 39
- Step 2: Establish and Enforce One Authoritative Source for Each Criterion 42
- Step 3: Reconcile Criteria 46
- Step 4: Set Up an Iterative Validation Framework 52

All Four Steps, 53 Times 54
The New One-Stop Shop Interfaces 57
Enforcement 60
How to Define Success 64
What We Can Do Now 68
Appendix A: The Bold Future of Integrated Benefits 70
Appendix B: What Success Looks Like 73
Appendix C: Open Questions 76
Appendix D: Citations for the 30 Years of One-Stop Shops Table 78
Appendix E: Citations for the What's Actually Worked Table 80

EXECUTIVE SUMMARY

Imagine this scenario:

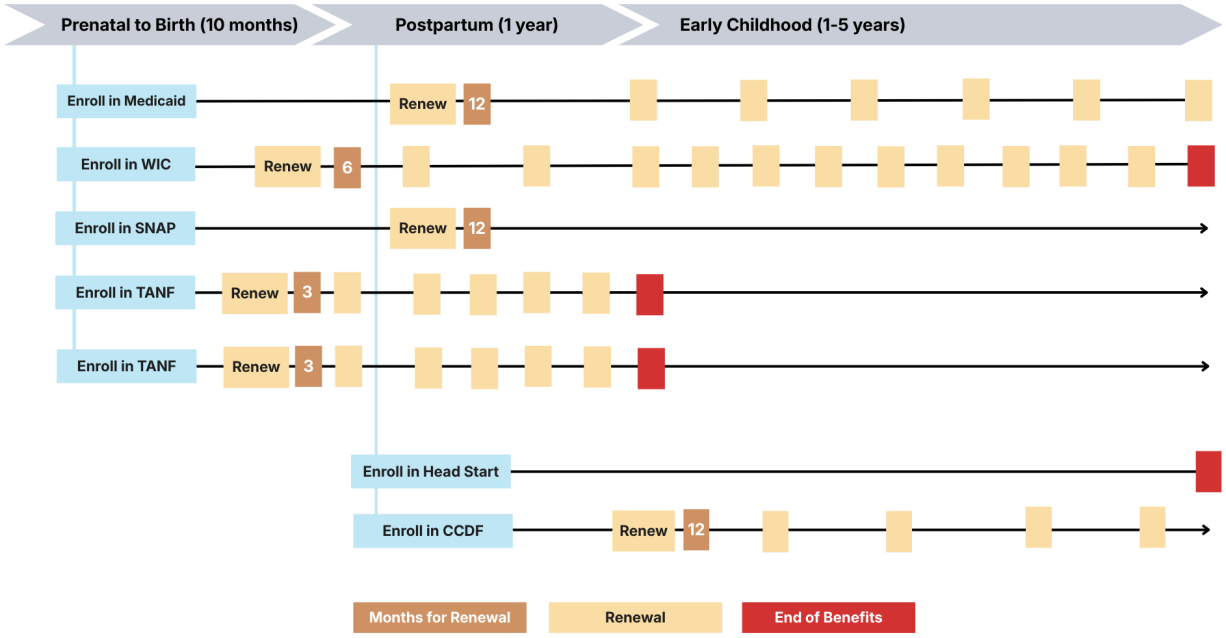
> Mirabel loses her job on Tuesday. By the end of the month she needs to file for unemployment, get food assistance, find new health coverage, cover an unexpectedly high heating bill, and figure out how to sign her kids up for free school lunch. This means at least five applications across five different agencies, all asking her approximately (but not exactly) the same questions, and none sharing her answers.

Despite claims to the contrary, an evil guy on the 12th floor did not purposely design this byzantine maze to punish Mirabel. Rather, no one designed the overall safety net *at all*. Every program only sees inside and only cares about its own silo. This "safety net" was created over more than 200 years of appropriations committees, agencies, federal accounts, programs, levels of government, and CIO dominions.

When Amira (one of this paper's co-authors) was the head of customer experience at the Office of Management and Budget (OMB), she interviewed a mother of a child with Down syndrome in 2018 as part of one of the first "life experience" pilots.[1] That mother described her reality: "*In America, you have to weave your own safety net.*" Later, while Amira was canvassing the country talking to low-income moms about navigating support in the first year of their child's life, another mother in North Carolina confided: *"I wish I had [support] when I was pregnant, in order to understand what all the programs mean. I still don't know all the things that are available to me."*[2]

Amira and her team mapped six benefit workstreams a low-income parent could be managing at any given moment for each child — while keeping many other balls afloat, like parenting, work, grocery shopping, and taking the kids to school. *(Page 7: Birth of a child life experience map from the Government Services Administration's Customer Experience Team.)*

[1] Federal Customer Experience Initiative Team, *Transitioning to Adulthood* (Washington, DC: Performance.gov, updated July 20, 2026).
[2] Federal Customer Experience Initiative Team, *Having a Child and Early Childhood* (Washington, DC: Performance.gov, updated September 29, 2026).



Source: "Birth of a Child" life experience journey maps available at https://www.performance.gov/cx/life-experiences/having-a-child-and-early-childhood/

The obvious response to this mess, which almost everyone arrives at on their own, is that there should be *one* place to go to get help — one primary application on one website that quickly and accurately assesses and delivers everything you're eligible for. It's what any reasonable person would design if they were starting from scratch.

When Marina (this paper's other co-author) became chief technology officer of the Department of Veterans Affairs in 2013, VA had an enormous number of websites (over 1,500!), each built by individual offices with their own budgets, eligibility criteria, claims processors, claims processing software, databases, logins, forms, contracts, and contractors. Veterans could not reasonably be expected to navigate it all, and they demonstrably did not.[3] Marina *herself* set out to achieve the goal of creating a one-stop shop.

User research told Marina's team repeatedly that not only did veterans not want to hear from the VA 80 times the moment they separate from the military, they would prefer not to hear from the VA *at all*. In response, the VA ran an ad campaign called "The Second Thing," highlighting veterans' dreams of the "first thing" they planned to do upon separation (like hugging their kids), then encouraging them to contact the VA as their second thing.[4]

[3] Anne Kainic, Albert Wong, and Jeff Maher, *Technically Speaking: How We Built a New Health Care Application for Veterans* (U.S. Digital Service on Medium, November 7, 2016).
[4] Alex Zamiar and Jonathan Richman, *U.S. Dept. of Veteran's Affairs: The Second Thing* (Squidbomb, 2014).

Addressing this concern by consolidating all VA benefits into a single front door — what started as Vets.gov, and later took over VA.gov — was among the things Marina most wanted to do. And she did it! Veterans today are generally able to apply for and manage all their federal veterans' benefits at VA.gov. The site was almost universally heralded as a huge success, even winning a 2018 Sammie award.[5] On the back of this success, VA was named lead of the Cross-Agency Priority (CAP) goal for customer experience for the entire federal government.[6]

Amira had her own experiences with one-stop shops. Conveniently able to access the details of agency budgets while at OMB, she started to pull the thread on Benefits.gov — a website that should, in theory, help solve the confusion of that mom in North Carolina. At the same time, there was a separate federal team, funded through annual appropriations, operating a website called USA.gov that was *also* supposed to be solving that confusion.

Benefits.gov aggregated information from across federal agencies, put links in one place, and had a short form to help filter benefits that might be relevant. USA.gov aggregated information from across federal agencies, put links in one place, and had a short form to help filter benefits that might be relevant. The government was paying for the same thing, twice.

OMB had already picked USA.gov as the winner for a "single front door." In 2004, OMB Memo M-05-04 required all federal websites to link to FirstGov.gov (which became USA.gov in 2007).[7] Government-wide digital analytics showed USA.gov had higher traffic than Benefits.gov; making things worse, the budget for Benefits.gov, operated by the Department of Labor, went almost exclusively to a single vendor. It took Amira three years to focus people's attention and generate concern over this; sunsetting Benefits.gov ultimately required an executive order in 2021, explicitly directing it to be consolidated into USA.gov.[8]

What we both ultimately learned from our VA.gov and USA.gov adventures — both *ridiculously* hard power struggles that continue today — is that building consolidated websites was the far easier, and far less impactful, part when it came to delivering for the American people.

Putting information or applications in one place does not inherently make programs (or the services they deliver) work together. Despite VA.gov, different VA benefits still define the same terms differently, still ask for the same proof separately, and still reach their determinations inside separate systems that don't communicate with one another. And that's inside *one* agency — consider the term "household" across tax agencies, Health and Human Services (HHS) funded programs, and Department of Agriculture (USDA) programs.

---

[5] U.S. Department of Veterans Affairs, *VA Employees among Federal Government Workers Honored with Sammies Awards* (VA News, October 3, 2018).
[6] Tim Hudak, *VA Releases Customer Experience Cookbook for Federal Agencies* (VA News, November 17, 2020).
[7] Clay Johnson III, *Policies for Federal Agency Public Websites*, M-05-04 (Washington, DC: Office of Management and Budget, December 17, 2004).
[8] Executive Order 14058, *Transforming Federal Customer Experience and Service Delivery To Rebuild Trust in Government*, 86 Fed. Reg. 71357 (December 16, 2021). The sun rose again in October 2026, when the Trump administration launched "America.gov."

VA.gov is a good place to start an application. But what happens after someone presses “submit” remains largely unchanged, and the underlying broken parts are now harder to see, because the site is functional and beautifully designed. For the applicant, this might actually be worse than the old mess. Back when they used to get stuck at the first step, at least they *knew* they were stuck; now they get stuck six weeks later, deep inside a process they thought they'd already finished, with no idea which of four offices has their file.

Recently, Amira’s current team researched whether integrating more benefits forms improved the “efficiency” of individual states’ programmatic administration spending. It did not — which means that while we may be saving people time in sharing their information with the government (which is good!), we are still not improving costs or making gains on the back end, because we are *not integrating the underlying work*. Every "modernization" that launches a website without changing any processes underneath teaches Congress and the public the same lesson: No matter how much money government spends, nothing gets better.

This is the sort of ultimately counterproductive work Jen Pahlka warned us about: Are we papering (or portaling?) over complexity, extending the life of a broken existing system that actually needs to die if we’re to achieve the future we want?[9]

A new website is the wrong goal in our pursuit of a simplified experience quickly getting what you need in a moment of financial insecurity. Instead, we need an intentionally designed, simplified, and authoritative government eligibility framework for public benefits.

This paper lays out the history of work on integrated benefits delivery. It then compiles insights from interviews with over 50 leaders responsible for integrated benefits delivery efforts at both the state and federal levels about why we keep failing and what we need to do instead. Spoiler: The problem is not incompetence or lack of trying. Instead, it boils down to control, data sharing, federalism, and the laws of cybernetics.

A cohesive eligibility and benefits delivery framework is what people are actually asking for when they ask for a single website. They’re really describing a user's desire to deal with government once, no matter what they need. This is entirely achievable, but only by doing work that looks nothing like building a website — and work that we haven’t really done before, let alone with a powerful mandate.

That gap between building a new website and doing the work to reconcile the eligibility criteria underneath is bigger than just benefits. The United States is very good at inventing technology yet has no systematic way of turning it into changed institutions and measurable public value. Benefits delivery is simply the clearest case study of what such a capability would have to do.

This paper lays out the task ahead, which consists of four parts:

[9] Jennifer Pahlka, *A Three Horizons Framework for Government Reform* (Eating Policy, April 24, 2026).

1. **Consistent, documented eligibility criteria.** Clearly map the objective eligibility criteria for every service the primary benefits agencies administer (beginning with at least SSA, USDA, HHS, HUD, DOL, IRS, and VA, and probably SBA, FEMA, and USCIS), in the same machine-readable format, so multiple benefits' criteria can be compared side by side. Each requirement should be cited as to whether it is in statute, a rule or regulation, federal guidance, state guidance, somewhere else, or "unknown." State-administered federal programs need to be documented at every lower layer.[10]
2. **Establish a single, government-run, authoritative source for each criterion**. Any criterion that is shared across benefits programs (like "income") requires a single, automated validation source. Once that exists, government (federal, state, local) *must use,* and may *only use* that source. This source must be held to (and resourced for) high standards of usability, availability, and accuracy.
3. **Reconciliation.** Where different programs use the same word to mean different things, we do the work to make each criterion a question the one shared authoritative source can answer.[11] This does not require every program to use the same threshold, only that "income" always means something the single, authoritative income service can answer accurately. Some of this will be easily agreed upon, and some will generate decades of debate. Might as well start now.
4. **Validation.** Every agency maintains a set of test scenarios that any system implementing its rules has to pass, and keeps these tests current. When a system gets something wrong, either the criteria or the tests get updated. Empowered caseworkers, within the agencies, who can exercise judgment and make binding decisions, are responsible for all scenarios that cannot be addressed by automation.

Do those four things and something new becomes possible: Any third party — an AI agent, a company, a nonprofit — can serve as an interface that can accurately assess eligibility across federal and state programs, and submit all necessary data and validations on a person's behalf across any number of programs.

At that point, you inherently have a one-stop shop. Maybe you have a million one-stop shops, with every American having their own personal government shop in their preferred font size and color scheme. While a government-run option may still be desirable, government stops having to be the primary provider and designer of all user interfaces — a role it's consistently struggled with — and instead focuses on the things only the government can do: set and affirm specific eligibility criteria, and deliver the subsequent benefits.

---

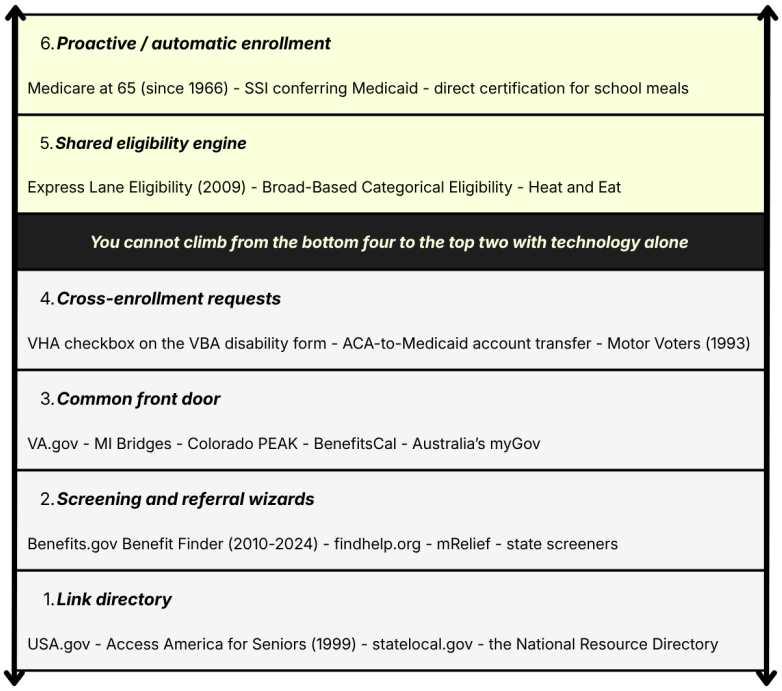
[10] Ariel Kennan, Lisa Singh, Bianca Dammholz, Jason Yi, and Keya Sengupta, *Exploring Rules Communication: Moving Beyond Static Documents to Standardized Code for U.S. Public Benefits Programs* (Washington, DC: Beeck Center for Social Impact + Innovation and Massive Data Institute, updated September 4, 2025).

[11] Arabella Crawford and Christopher Whitaker, *Standard Definitions: Reducing Administrative Burden by Ensuring Interoperability* (Washington, DC: Searchlight Institute, July 27, 2026).

The Six Rungs of "Integrated Benefits"

The six rungs form a rough ladder that takes us from a website that points at forms up to a government that dispenses benefits before it's even prompted.

The six rungs of "integrated benefits"



Rung 1: A link directory. One website that organizes the links, forms, and information across multiple benefits, and potentially multiple levels of government (or even public and private sector resources). Users may find what they need faster, but ultimately are directed to individual programs' websites and forms.

Examples:

- USA.gov
- Access America for Seniors (1999)[12]
- The U.S. State and Local Gateway (a specialized web portal at statelocal.gov launched in the late 1990s by the National Partnership for Reinventing Government)[13]
- The National Resource Directory (jointly maintained by DoD, VA, and Labor for wounded warriors).[14]

[12] The White House, *Technology Access and Innovation* (Washington, DC: The White House, June 2000).
[13] National Partnership for Reinventing Government, *History of the National Partnership for Reinventing Government: Changing the Culture 1997–2001* (Washington, DC: National Partnership for Reinventing Government).
[14] National Resource Directory, *National Resource Directory* (Washington, DC: National Resource Directory).

Rung 2: Screening and referral tools. The system collects information from applicants and surfaces (without authority) what benefits they *might* qualify for, then directs them to external sites to apply.

Examples:

- The Benefit Finder on Benefits.gov (2010 until the site closed in 2024) and its successor on USA.gov
- Screeners built into Michigan's MI Bridges and Colorado's PEAK
- ACCESS NYC
- MyFriendBen
- Findhelp.org
- mRelief

Rung 3: A common front door. One portal with a consistent look and feel, and one login that grants access to multiple applications. Users may be able to see the status of submitted applications and reuse some stored data elements, like mailing address, to prepopulate new applications. The site may be able to suggest applications for additional benefits and generate notifications and reminders.

Examples:

- VA.gov
- Michigan's MI Bridges
- Colorado PEAK
- California's BenefitsCal
- Frontdoor Benefits
- Mirza

Rung 4: Cross-enrollment requests. A checkbox on one program's application opts applicants into another benefit. This is typically only for first-time applicants, as recertification timelines generally don't line up. (Another opportunity for reconciliation!) One fewer application is one fewer application, but this enhancement is mostly cosmetic, because the checkbox just copies your file over to a second agency, which runs its own process. This only works for programs with *truly* overlapping application information, which is uncommon.

Examples:

- The account transfer from the ACA marketplace application to state Medicaid and CHIP agencies
- The National Voter Registration Act of 1993, which requires states to offer voter registration at the DMV and at public assistance offices

Rung 5: Shared eligibility engine. One application and verification used across multiple programs, and potentially reused again at recertification. This rung includes “deeming,” where one program formally accepts another program's finding without reverification.

Examples:

- Express Lane Eligibility, created by CHIPRA in 2009: Medicaid and CHIP agencies may accept another program's finding on income, household size, or citizenship instead of reverifying it. States have used SNAP, TANF, WIC, and school lunch as the express lane agency. The Medicaid agency still applies its own threshold to the borrowed findings, so this shares verification without sharing exact criteria.
- The Combined Application Projects, an SSA and FNA demonstration in roughly nine states, does something similar for elderly and disabled SSI recipients
- Broad-Based Categorical Eligibility lets SNAP treat receipt of a TANF-funded benefit as proof of SNAP eligibility, bypassing its own asset test — an approach analysts at AEI have criticized as so loose it defeats the purpose of having separate criteria, as a “TANF-funded benefit” can be receiving a brochure.
- Heat and Eat: receipt of a LIHEAP payment in participating states automatically qualifies a SNAP household for the standard utility allowance, raising its benefit without any new verification
- The new disaster assistance application launched in 2023 by FEMA and SBA

Rung 6: Proactive / automatic enrollment. The government enrolls or reenrolls an individual in a benefit without requiring action on the part of that individual. This might use population-level eligibility, or act on already available data. This rung is the least likely to exist because it requires the most substantive lift. In service design, getting what you need via “no experience” is generally the best possible experience.

Examples:

- Ex parte Medicaid renewals, where the agency redetermines eligibility from data it already holds
- Most states automatically enroll SSI recipients in Medicaid.
- Children in SNAP or TANF households are certified for free meals without anyone filling out an application. Summer EBT works the same way: Eligible children can be automatically enrolled in SUN Bucks through SNAP, TANF, FDPIR, or other qualifying programs.
- The pandemic child tax credit and the Economic Impact Payments were both delivered this way at national scale.
- The largest example has worked quietly since 1966: Anyone already receiving Social Security when they turn 65 is automatically enrolled in Medicare Parts A

and B. SSA already knows their age, identity, and work history, so it has everything needed for enrollment.

- Enrollment in a Medicare Savings Program automatically deems a person eligible for the Part D Low-Income Subsidy, with no separate application.
- In Canada, filing individual tax returns enables the federal government to automatically assess your eligibility for several federal and provincial benefit and credit payments using your reported net income.

Table: The Six Rungs of Integration

Rung	What it requires	What it delivers	What it cannot fix	Who has to agree
6: Proactive / automatic enrollment (Top rung)	Redesign of existing processes; new authorities and data sources; ability to act on data collected rather than wait for "opt-in"	Benefits with no application at all; streamlined recertifications; elimination of "churn" and processing work; higher enrollment	Cannot require data it can't access	Congress
5: Shared eligibility engine	Definitional reconciliation; regulatory and statutory authorities to empower centralized determination; agreement on who owns/runs it; ability of central authority to adjudicate and enroll rather than just pass data off	Sharing data once for multiple programs; increased efficiency on back end for reduction of adjudications and reuse of verifications	Individuals still have to apply; engine itself can only surface, not fix, eligibility conflicts	Every owning agency, on every shared term and on how they will integrate action; generally, Congress too
4: Cross-enrollment	Precisely overlapping data / criteria; data-sharing agreements; light technical work to route applications to adjudicating agencies	A warm handoff + first action for another benefit (apply); saves applicants time; some error reduction (as application data is always the same)	Everything downstream of "apply"; disjointed eligibility criteria; not delivering internal efficiency gains; cannot streamline reenrollment between programs because data is now forked	Statutory or regulatory guidance owners to accept others' data
3: Common front door	A shared identity layer; agreement on intake; awareness of application processing status	One login, one status view; some prepopulation; two-way communication	Eligibility disagreements; benefit cliffs; time to determination	Participating agencies' intake and application processing functions
2: Screening and referral tools	Published eligibility criteria; a very basic information collection form (does not need to store or validate)	Some personalization of what benefits are relevant for you	Whether the person can finish the application; whether the screen was right	Whichever agencies publish criteria
1: Link directory (Bottom rung)	Very little or no effort from programmatic operators; can be done by a third party; content maintenance; no policy changes	Findability	Everything past the basic "learning" stage of benefits access	Whoever owns the website

Table: Searchlight Institute

Most efforts have gone into climbing the bottom rungs

The bottom four rungs are interface work. A directory, a front door, a screener quiz, and a cross-enrollment checkbox all sit on top of the existing eligibility machinery. They *can* change a person's experience, but none of them requires any program to make any changes. These projects are popular because they are achievable within a single term, they demo well, and they require little buy-in and no systemic change.

But the two at the top instigate action. A shared eligibility engine requires programs to accept one another's findings. Proactive enrollment requires an agency to access facts sufficient to make eligibility decisions, and exercise the authority to act on them. These are the rungs that create meaningful change, and change at this scale is hard — which is why these examples are much rarer. Notice that none of what makes them hard is technical. The top rungs change who decides, who is accountable when the decision is wrong, what caseworkers do all day, and which legal authority is being exercised. That is institutional change, with software underneath it, and it's why every project that treated the top rungs as a technology deployment ended up in the failure table.

You cannot climb from the bottom four to the top two

This taxonomy should not be mistaken for a maturity model. Build a good directory, then a good front door, then a good screener, and eventually you will arrive at a shared eligibility engine, right? Once the other agencies see your beautiful website, read the work of your talented plain language writers, and see your user count soar, they'll be lining up to hand over their forms, websites, and eligibility processes.

Except: *They will not*. Climbing the bottom four rungs well, doing each step better, for longer, with more money and bigger teams, will *never* enable you to reach the two at the top. A screener that has been refined for 22 years is still a screener, because nothing about refining it obliges any program to so much as agree on a single word with any other program. The rungs are not on the same ladder; they are not in the same universe. The bottom four are software and design problems. The top two are policy problems disguised as software, and the only way to reach them is to do the policy work. **Concretely, that policy work is: writing down what every program's rules actually are, in one format; reconciling the rules that differ for no reason anyone can defend; naming one source the government will trust for each fact; and proving that the systems apply the rules the law actually wrote.**

The rest of this paper is about those four tasks, and about where they lead: a safety net where most determinations happen automatically, from facts the government already holds, and where caseworkers have the time and capacity to genuinely help those who need it most.

30 YEARS AT THE BOTTOM OF THE LADDER: A HISTORY OF ONE-STOP SHOPS

Every administration since 1993 has built at least one portal.[15] The table below lays out the history: Nearly everything on the bottom four rungs shipped and is still running under some name; every attempt to jump straight to the fifth rung failed, expensively; and 30 years of effort at the bottom didn't produce a single step up.

Table: 30 Years of One-Stop Shops

Year	Effort	Rung	What happened
1993	National Performance Review takes up electronic government	—	The federal government's first digital government, one-stop shop goal. Published as Access America in 1997.
1999	Access America for Seniors (SSA)	1	Change your address with USPS, compare nursing homes, request a Social Security statement, and download IRS forms. All links on one page.
2000	FirstGov.gov	1	President Clinton ordered a launch within 90 days of his announcement; it went live 87 days later, during the first-ever webcast from the Oval Office. A government-wide search engine spanning 47 million pages. Renamed USA.gov in 2007.
2002	E-Government Act's "Federal Internet Portal"	1	Section 204 directed OMB and GSA to create a portal for "Internet-based Government information and services directed to key groups, including citizens, business, and other governments, and integrated according to function or topic rather than separated according to the boundaries of agency jurisdiction."
2002	GovBenefits.gov	1	One of 24 initiatives in President Bush's E-Government Strategy. DOL was put in charge, with 10 agencies contributing content.
2004	Colorado CBMS	5	Failed; $100M to try to replace six legacy systems with one eligibility engine. State auditors called it a disaster. Researchers titled their case study of it Seven Years of Failure.
2010	GovBenefits.gov rebranded as Benefits.gov	2	Adds "Benefit Finder," a screener. Grows to 17 agency partners with information on more than 1,000 programs.
2012	MyGov / MyUSA (Presidential Innovation Fellows)	3	A prototype to reorganize federal websites "by citizen needs rather than the complicated organizational chart of the federal bureaucracy." Never launched.
2013–15	Cover Oregon	5	Failed; After roughly $240M, it never enrolled a single person. Dissolved by legislation in March 2015.
2015	Vets.gov (later VA.gov)	3	Over 1,500 VA websites consolidated into one. Still alive and kicking.
2016	Rhode Island UHIP	5	Failed multiple times, generating a public apology from Deloitte.
2021	Executive Order 14058	—	Directs agencies to improve customer experience and designates USA.gov the federal "front door."
2024	Benefits.gov shut down	—	After 22 years, finally forced to consolidate into USA.gov. Celebrations were had.
2026	OMB M-26-18, "Unlocking More Integrated Experiences"	3	Directs agencies to use information, with consent, to prepopulate forms across services: "greatly reducing the number of times that users need to provide the same information."
2026	America.gov launches	3	An AI chat interface, which pulls significant data from USA.gov, launches offering additional search functionality.

Table: Searchlight Institute • Source: See Appendix D for citations.

[15] Erie Meyer, *Should I Build a Government Portal?* (Medium, November 15, 2020).

Vice President Al Gore's National Performance Review took up electronic government in 1993, publishing *Access America* in 1997.[16] It produced FirstGov.gov, along with statelocal.gov for state and local employees. In February 1999 it launched Access America for Seniors, hosted by the Social Security Administration, which let older Americans change their address with the U.S. Postal Service, compare nursing homes, request a Social Security statement, and download IRS forms in one place. FirstGov became USA.gov, which is still with us, having been reinvented a number of times.

In April 2002, the Bush administration launched GovBenefits.gov as one of 24 initiatives in the President's E-Government Strategy, with the Department of Labor designated as lead and 10 agencies contributing content.[17] It grew to 17 agency partners with information on more than a thousand programs. It was rebranded as Benefits.gov in 2010, when it added a "Benefit Finder" (a Rung 2 screener) and later a chatbot.

In September 2024 Benefits.gov was shut down and absorbed into USA.gov, as directed by Executive Order 14058.[18]

In a different vein, in 2012, Marina's colleagues in the first class of Presidential Innovation Fellows took on a project called MyGov, which was going to "create a prototype of a streamlined online system enabling citizens to easily access the information and services from across the Federal Government" and reorganize federal websites by citizen needs rather than the complicated organizational chart of the federal bureaucracy. It was later renamed MyUSA. You have not used it, because it never saw the light of day.

In 2015, a team from GSA did an extensive research effort to improve USA.gov.[19] Again, the user needs they gathered could not be addressed by a better website; solutions required the hard work of cross-agency data and process integration.

In 2022, the "Facing a Financial Shock" team (which sits within the life experiences effort that Amira led as part of the customer experience executive order) started to stretch for higher rungs. It took a White House principals' meeting to agree to create a mandate for the work, and extensive central coordination to ultimately result in the deputy secretaries of Agriculture, Health and Human Services, Labor, Treasury, Housing and Urban Development, the administrator of the General Services Administration, the commissioner of the Social Security Administration, and the deputy director for Management signing a charter agreeing to the goal that Americans seeking crucial services be able to:[20]

---

[16] Al Gore, *Access America Introduction by Vice President Al Gore* (Washington, DC: The White House, February 1997).
[17] Office of Management and Budget, *Access to 30 New Programs Drives Success of GovBenefits: Web Site Assists Over 13,000 Citizens a Day*, 2002-44 (Washington, DC: Office of Management and Budget, July 16, 2002).
[18] Executive Order 14058, *Transforming Federal Customer Experience and Service Delivery To Rebuild Trust in Government*, 86 Fed. Reg. 71357 (December 16, 2021).
[19] USAGov, *Informing the Future of the Federal Front Door* (Washington, DC: USAGov, updated February 2, 2026).
[20] President's Management Council, *Life Experience Designation Charter: Facing a Financial Shock and Becoming Newly Eligible for Critical Supports* (Washington, DC: Performance.gov, 2022).

- Apply in 20 minutes;
- Enroll in 24 hours;
- Only have to share their information once; and
- Receive services within a week.

This “North Star” vision helped kick off a State Benefits Leadership cohort that opened communication on tactical work to achieve this outcome between administrators at the federal and state levels.[21] The federal mandate and vision helped affirm creative work some states were already doing, de-risk work for those that were unsure, and provide technical assistance for those that didn’t know where to start. But, with no one empowered to actually make any of this happen, it was unable to be more than a “culture change” effort.

States have also made many runs at climbing the ladder. Minnesota's MNbenefits.mn.gov, built with Code for America and New America, lets residents enter their information once to apply for as many as nine safety net programs in under 15 minutes. Michigan has MI Bridges. Maryland's Benefits One Application covers five programs. As of 2025, Code for America counts 35 states that offer an integrated application spanning at least three programs.[22]

These are real accomplishments, and we are not dismissing them — Marina started Vets.gov, after all. **But these state- and federal-level projects all hit the same limits: They couldn’t influence anything beyond the front door. Sure, people could apply for nine benefits in 15 minutes, but they still faced long processing times (sometimes *years*), rejections when they should have been approved, and other obstacles**.

Over 30 years of federal and state effort, and a lot of money, has gone into the bottom four rungs. And pretty much everything on the bottom four rungs shipped. Some of it even shipped quite fast; FirstGov went from presidential announcement to live website in 87 days, which is a delivery record most federal programs today would envy. Today, almost all of the programs are still running in some form, under some name, having been rebranded or absorbed at least once.

In contrast, every single attempt at Rung 5 failed. Colorado spent $100 million and got a system its own auditors called a disaster; a subsequent federal study identified the specific problem as "difficulties updating CBMS's rules engine to accommodate different program requirements." Oregon spent $240 million and enrolled nobody. Rhode Island went live with UHIP and it promptly collapsed. Tennessee's TEDS, Alaska's eligibility backlog, and a myriad of additional projects belong on this list, too. Every postmortem story is the same: Each set out to put eligibility determination for many programs in a single place, and each discovered that the individual programs did not agree on what they were determining, would not cede power, and would undermine the consolidated system until it died.

---

[21] Cassandra Madison, Tyonka Perkins Rimawi, and Tim Shaw, *North Stars in a Moment of Constraint* (Washington, DC: Aspen Institute, September 9, 2026).

[22] Code for America, *The Benefits Enrollment Field Guide* (San Francisco, CA: Code for America, updated September 2025).

Thirty years later, the front doors keep getting rebuilt with newer technology, renamed, and folded into the *next* front door, while the rules underneath stay exactly where they were. That's what has to change.

EIGHT LESSONS FROM 30 YEARS OF INTEGRATED BENEFITS DELIVERY FAILURES

Interviews with more than 50 leaders of state and federal integrated benefits efforts surfaced the same handful of obstacles, regardless of era, party, or technology stack — as well as valuable insights into what *has* worked. For any federal (or state) effort to be successful in the future, both the shortfalls and successes from these past efforts must be learned from.

1. Nobody Gives Up Power

"When you're trying to diagnose why a system is failing, you need to consider both the larger system in which it's embedded and the organization within its operations." — Dan Davies, *The Unaccountability Machine*

The first and largest obstacle is that no program, agency, office, or leader is going to voluntarily hand over control of its eligibility rules or benefits processing — and there may be actual laws tying agencies' hands as well. Yet, consolidating power is exactly what a shared determination engine requires, which is why every one of those engines is in the failure table above. This span of control exists across forms, instructions, and even website colors.

While we may have wanted to purchase hexes against them on Etsy, the officials who resist giving up control are not irrational. These officials' statutory obligation is to administer their program correctly. Accepting another agency's definition or website or form or uptime means accepting responsibility for outcomes they no longer control, all while knowing that no oversight body will accept "but we delegated this to the integrated benefits portal!" as an excuse. If they were simply being obstinate, you could replace them with more cooperative people and get a different result. But that won't work: Marina, for example, would never, ever, ever have given up so much as her desktop printer to be administered by GSA, let alone handed over VA.gov. (Amira did, however, find GSA to be a superior source of pens and paper, as their office supply procurement was, shall we say, a little better than OMB's.)

Consider that from 2002 until 2024, GovBenefits.gov (later Benefits.gov) was run by the Department of Labor. DOL does not administer SNAP, Medicaid, CHIP, TANF, WIC, or LIHEAP. It ended up as the lead agency because that's what the Bush E-Government Strategy decided in 2002; as best we can tell from digging through records, this was because it was originally DOL Chief Information Officer Patrick Pizzella's idea.

GSA, meanwhile, runs shared services for the federal government as its *actual job*, and has been operating the *other* federal front door (FirstGov, then USA.gov) since 2000. It may seem obvious to consolidate both front doors into one, and for that consolidated portal to live at GSA. But DOL was not going to let it go so easily. Once an agency has a thing, it does not historically let go.

Benefits.gov finally shut down in September 2024, after a multiyear, painstaking effort during which Amira worked to align stakeholders and support across White House offices and agencies to require the consolidation in Executive Order 14058. The EO reaffirmed USA.gov as "the front

door to government information and services." (Note: Much of Benefits.gov's traffic originated on USA.gov, which adds up to ... two stops, not one).

In the same 22-year period, the number of conflicting federal definitions of "income" that were reconciled via executive order (or any other mechanism) was zero.[23] If moving a domain name takes two decades and an executive order, consider what it may take to make USDA and HHS agree on who counts as a household, let alone how often it needs to be verified, and by whom.

DOL has a storied history of domain squatting. It also owns Veterans.gov. As CTO of the VA, Marina spent over a year trying to get the domain, which DOL was not even using at the time. Multiple White House meetings tried and failed to resolve the extremely complicated question of whether or not DOL should be allowed to squat on the domain name. With the president's convening authority, a government-wide digital strategy, and every reasonable argument on her side — Marina lost. Instead, DOL secretly launched a poorly designed website that analytics data shows no one but bots visits, *the same day* and *at the same exact event* where VA launched Vets.gov, after demanding Marina trek to their office for weekly deskside demos of Vets.gov every week up to launch.

As these two examples show, we challenge anyone who believes in getting agencies to voluntarily integrate their benefits to first try getting one of them to voluntarily give up *anything*.

Anthony Downs identified this phenomenon in *Inside Bureaucracy*: The greater the effort a top-level official makes to control the behavior of subordinate officials, the greater the effort those subordinates make to evade or counteract that control.[24] Any plan that depends on an agency to voluntarily surrender control will lose. But that means a plan that does not ... can win. **That is the difference between a memo and a statute. An agency can slow-walk a memo indefinitely; it can't slow-walk an appropriation that doesn't arrive, or a law that moves the authority somewhere else. The counter-control Downs describes is strongest where the control is weakest, which is why most of what we propose later runs on appropriations and statute rather than memorandums of understanding (MOUs)**.

2. Programs Will Not Share Data or Trust Shared Data

Agencies collect and store the same information about individuals repeatedly because the alternative requires sharing data and trusting shared data. That's hard for at least four separate reasons that are often lumped together:

---

[23] U.S. Government Accountability Office, *Federal Low-Income Programs: Eligibility and Benefits Differ for Selected Programs Due to Complex and Varied Rules*, GAO-17-558 (Washington, DC: U.S. Government Accountability Office, June 29, 2017).

[24] Anthony Downs, *Inside Bureaucracy* (Santa Monica, CA: RAND Corporation, 1967).

1. Data-sharing agreements are slow, narrow, revocable, and negotiated between agencies one data set at a time. Every new program in a shared system multiplies the number of agreements needed.

2. Even with an agreement in hand, there's the realization that "income" definitions are not standardized between agencies, so sharing isn't that fruitful.

3. Most eligibility systems were built to take in data from applicants, not to send or receive it with other agencies. They may not be able to incorporate the data in a meaningful way that improves business processes.

4. Accepting data that was "not collected here" means trusting someone else's process. On the flip side, sending your data elsewhere means remaining accountable for what happens to it downstream.

SBA for years has had issues with how GSA has made determinations on its behalf, even becoming the subject of an IG investigation.[25] Having been burned before, SBA was not irrationally skeptical of data sharing when it came to collaborating on disaster assistance with FEMA; that cross-agency data sharing was one of the biggest blockers to developing a streamlined individual assistance application. Yet disaster assistance is now a Rung 5 success story: The “Recovering From a Disaster” life experience effort helped coordinate a cross-agency steering committee that streamlined the policy for applying for SBA assistance prior to applying for individual assistance.[26] Success is possible.

We are also not going to pretend in 2026 that a carefully negotiated data-sharing agreement stops a sufficiently motivated actor from unapproved access or data reuse. On April 7, 2025, the IRS signed an MOU with ICE to share taxpayer information for immigration enforcement.[27] IRS lawyers advised that the deal likely violated privacy law. The acting commissioner resigned over it. That June, ICE requested 1.28 million taxpayer records; by August, the IRS had already sent them tens of thousands of records, including home addresses.[28]

In November, a federal judge barred the IRS from sharing addresses with ICE. In the litigation that followed, the IRS' own chief risk and control officer acknowledged that the agency had improperly shared that data, violating not just federal law but the terms of its own MOU. The court determined that the IRS had broken federal law approximately 42,695 times.[29]

---

[25] Jason Miller, *SBA Finds Fault with GSA’s Strategic Sourcing Analysis for Office Supplies* (Federal News Network, April 2014).; Mark Rockwell, *IG: GSA System Erred on Small Business Designation* (Nextgov/FCW, September 17, 2020).

[26] Federal Customer Experience Initiative Team, *Recovering from a Disaster* (Washington, DC: Performance.gov, updated December 2024).

[27] Justin C. Chung, *Centro de Trabajadores Unidos v. Bessent: D.C. Circuit Considers IRS-ICE Information-Sharing Agreement*, LSB11359 (Washington, DC: Congressional Research Service, September 4, 2025).

[28] Fatima Hussein, *Data of Thousands of Taxpayers Wrongly Shared with DHS, Court Filing Says* (PBS News, February 12, 2026).

[29] Jeff Stein, *Judge: IRS Broke Law “Approximately 42,695 Times” in Giving DHS Data* (The Washington Post, February 26, 2026).

So, will a sufficiently detailed data-sharing agreement protect against misuse? No. That agreement was negotiated by lawyers who thought it was illegal, signed over the objection of the agency head, and then breached tens of thousands of times by the agency that wrote it. A data-sharing agreement is a promise about future behavior, and the trouble with that is the person who made that promise is not always the one responsible for keeping it.

Every state benefits agency watched this happen. Now when a state hesitates to hand data to a federal partner, or to accept a data-matching arrangement with another agency, that hesitation is viewed sympathetically rather than as obstructionist. This makes the job of expanding data sharing for the purpose of integrated eligibility that much harder.

Let's keep in mind that proactive enrollment does not always require data sharing. SSA automatically enrolling its recipients in Medicare when they turn 65 works because SSA already holds every fact it needs to do that. Ex parte Medicaid renewal — when a state uses administrative data it already has to renew enrollees without requiring an individual to submit additional documentation — works the same way.[30] But the moment delivery has to cross a program boundary — like using Department of Education data to streamline access to Medicaid — objections arise.

A solution will require hard, clear, new approaches to data collection, storage, sharing, and consequences for misuse (including safe harbor protections for sharing and reusing data), which will help rebuild trust. We take that up in a separate, future paper. For this one, **the takeaway is that integrated delivery depends on data moving between programs, and today it mostly doesn't: for legal reasons, because the definitions don't line up, because the systems were built to take data in from applicants rather than from each other, and because nobody is accountable for delays or inability to access benefits**.

3. A Shared Eligibility Engine Is Really *Two* Problems

Many people deep in this work have pointed out that shared eligibility gets treated as a single hard thing when it is actually two different things, with potentially two different answers.

Declaring who is eligible is not the same thing as declaring what is needed to prove you are eligible. Part of what *looks* like shared determination is shared verification of a raw fact. Citizenship, date of death, incarceration status, and veteran status are facts, not subjective decisions.[31] If SNAP has confirmed citizenship, there is no reason Medicaid should have to check again. This part is tractable, and much of what we recommend later is built on it.

[30] Pamela Herd, Eric R. Giannella, Jeremy Barofsky, Luke Farrell, and Donald Moynihan, *Interventions To Automate Medicaid Renewals Reduce Procedural Denials And Increase Coverage*, Health Affairs 44, no. 11 (2025): 1336–1343.
[31] Yes, the question of whether someone "is" a veteran can be complicated, including the character of discharge, and whether that character of discharge can be corrected. See: U.S. Department of War, *Don't Ask, Don't Tell Resources*.

There's one catch: Sharing a verification means agreeing on how sure you need to be that it's accurate. Residency is a legal requirement for both voter registration and purchasing a gun; plenty of people want it proven lightly for one and heavily for the other (often the same people, in opposite directions). Benefits programs will have exactly this fight, with different programs requiring different levels of assurance and using different sources of validation.

The second problem is what's *not* so easily shareable, because it uses the same word but has an entirely different purpose. A workers' compensation disability exam and an SSI disability determination are not two measurements of one fact; they are different legal questions that happen to share a word. A VA disability rating measures service-connected impairment and says nothing about whether a veteran can work, while SSI applies a five-step test of inability to work. Then there's unemployment insurance, which asks whether you are "able and available" for work, which is in some ways the opposite question. No amount of reconciliation fixes this, because the programs are at cross-purposes. **The eligibility criterion is the program's substantive policy, and should stay with the agency that owns it**.

But there's a different case that *is* fixable. For example, SNAP and Medicaid both exempt some American Indians and Alaska Natives from work requirements, for the same underlying purpose, yet use standards that are *just* different enough to prevent either program from reusing the other's finding.[32] Same purpose, different criteria, and no good policy reason for the delta. This should be on the reconciliation short list.

A possible rule of thumb: If two programs are asking for similar information for the same *purpose*, we should aim to reconcile. If they are asking for *different* purposes, even if the words are the same, we should not.

If SNAP says a person is a citizen and Medicaid says they are not, somebody made an error. That is a fact, so it belongs in the shareable bucket. If VA rates a veteran at 70% disabled and SSI finds them not disabled, both agencies can be entirely correct, because they are asking entirely different questions; that mismatch doesn't require addressing (unless we want to debate the core purposes of these programs).

One more distinction on the shareable portion that people pointed out: Some facts are immutable, while some change over time. Date of birth does not change. However, income, household composition, address, and employment status are true only as of a given date. Those need a timestamp and an expiration (or a live query of the authoritative source).

4. There Are *Really* 53 SNAP Programs

[32] Jennifer Wagner, Ed Bolen, Alicia Huguelet, and Dottie Rosenbaum, *Coordinating Medicaid and SNAP Work Requirements to Streamline Determinations* (Washington, DC: Center on Budget and Policy Priorities, updated September 2, 2026).

State-administered (or county-administered) federal programs are not all the same. Sometimes intentionally, and sometimes not, states vary in big and small ways. **Any sort of reconciliation requires understanding every single jurisdiction's rules, not only the federal layer**.

For example, SNAP is a single federal program administered by 50 states, the District of Columbia, Guam, and the U.S. Virgin Islands. As Rebecca Piazza, who previously worked on benefits delivery at 18F, Nava, and FNA, puts it: "There are *really* 53 SNAP programs. And you might say SNAP has it easy: There are 88 WIC programs."[33]

Multiply that by every program a federal portal would have to cover. **The complexity in rules and definitions — seen across these iterations of each program — is unmanageable, and that is part of why integrated benefits portals for state-administered programs have not worked at scale.**

It's tempting to conclude that the answer is to let each state build its own integrated system. But the rules come in layers: federal statute, then federal regulation and guidance, then the options a state has chosen (the choices federal law lets it make about its own criteria), then the state's own statutes and regulations, then its policy manual, and then sometimes a county and/or field office practice layer. A state controls some of that, but not all. Minnesota can reconcile the SNAP options it chose with the Medicaid options it chose. It cannot reconcile SNAP's federal definition of household with Medicaid's, because neither of those is Minnesota's to change.

5. The Code Does Not Always Match the Law

On the topic of locally administered federal programs, there's a problem underneath all of this that gets almost no attention, which is that both intentionally and unintentionally, state- and county-run systems regularly deviate from their governing legislative details in practice.

Software code and policy slowly drift from statute, and when it comes to the legislative code or the benefit system code, the benefit system's rules always win, by virtue of being the actual facts on the ground.

During the 2023 Medicaid and CHIP unwinding, systems in 29 states were found to be evaluating renewals at the household level rather than the individual level required by federal rules, wrongly terminating coverage for hundreds of thousands of children and adults.[34] Tennessee's system reverted to giving postpartum mothers two months of coverage instead of the 12 they were entitled to, without anyone noticing — it only surfaced when a beneficiary lost coverage and pushed back. In *Cholvin v. Wisconsin Department of Health and Family Services*, a court found

[33] U.S. Department of Agriculture, Food and Nutrition Administration, *About WIC* (Alexandria, VA: USDA FNA, updated November 24, 2025).
[34] U.S. Government Accountability Office, *Medicaid: Federal Oversight of State Eligibility Redeterminations Should Reflect Lessons Learned after COVID-19*, GAO-24-106883 (Washington, DC: U.S. Government Accountability Office, July 18, 2024).

that an assessment system misinterpreted the law by dropping functional limitations from consideration, imposing an eligibility standard that was never actually a rule. In *Darjee v. Betlach*, Arizona's HEAPlus system was found to have been applying a five-year residency bar to refugees who were actually entitled to full coverage.

Many of these systems run on older, less familiar code bases maintained by skeleton crews. The capability to generate a complete list of system rules is mostly nonexistent. But the advent of artificial intelligence is a new opportunity to rapidly assess these systems' underlying logic and compare them to the legislative and regulatory language.

Michigan's MiDAS shows what it costs when integrated portals go badly. The system adjudicated unemployment fraud determinations with no human review, ran at roughly an 85% error rate, and wrongly accused approximately 40,000 people of fraud between 2013 and 2015. Innocent people had their wages garnished, lost their homes, and filed for bankruptcy.[35] The state defended the system right up until litigation ultimately stopped it; it settled for $20 million in 2022. That same year, Michigan hired Deloitte for $78 million to replace the underlying system. The project ran 14 months late and needed another $20 million, all while keeping the old system alive was costing $2.5 million a month. As of this writing, the new system's benefit functions are scheduled to go live in November 2026, more than a decade later.

Systems have been deciding cases with incorrect criteria since the advent of computing, and there is no cohesive mechanism to identify and correct inconsistencies. **Understanding the eligibility rules that are actually in play and aligning them to legislation could bring much greater consistency to programs, opening the door to more successful integration** — but if the reward for finding your own errors is a lawsuit, states won't look. Federal lawmakers need to create a safe harbor window in which a government can find and fix its own mistakes without fear of a surge of lawsuits related to improper program implementation, for some period of time. We come back to this in the Enforcement section.

6. Ashby's Law of Requisite Variety

Ashby's Law of Requisite Variety, a concept from cybernetics management theory, holds that a regulator must have at least as much variety (the number of responses available to respond) as the system it regulates. Only variety can absorb variety. This is another explanation for why the portal-first approach fails, independent of execution quality.

An application portal is a regulator. It sits between the public and a set of programs, meant to resolve their complexity into a single, coherent interaction. The "variety" on one side is the multiple definitions of income, the program-specific household rules, the use of the term

[35] Paul Egan, *Michigan Makes Amends for Falsely Accusing Thousands of Unemployment Fraud* (GovTech, August 15, 2017).

“disability” to mean at least three entirely different things, the 53 SNAP programs, and the state and local customizations underneath each of them.

If the portal has less variety than that, it cannot eliminate the difference. It can only relocate it. A portal with less variety looks like one income field where six programs need six answers, or one “household” question that two programs interpret differently. The variety resurfaces as an applicant who cannot understand why they qualified for one program and not another, an applicant who submits an incorrect application because they didn’t understand what was being asked, a caseworker reconciling by hand what the system could not, or an incorrect disqualification leaving behind a desperate person in need.

Ashby’s Law applies to caseworkers, too. A lot of the current worry about AI in benefits is that the systems will get too complex for a human in the loop to check. *But that's the system we already have now*. The national SNAP payment error rate has been over 10% for several years running, with trained human caseworkers making those decisions from a manual.

The lesson is that there’s work to do on both sides: create a more capable regulator (more options and flexibility in the user interfaces and in a caseworker’s toolkit) and a less complicated set of eligibility rules, criteria, and programs for it to regulate.

7. A Front Door Is Still a Pull, Not a Push

A “pull” is a benefit you have to get. A “push” is a benefit that comes automatically. A front door, however good, is still a pull.

VA.gov is, by any reasonable standard, the best-case outcome for a consolidated front door. Over a thousand websites became one, and the mobile app has millions of active users. And yet, veterans still do not always know or get the benefits they’re entitled to. Many still call VA, rely on veterans services organizations for navigation support, or simply miss out.

The National Academies' 2018 evaluation of VA mental health services found that among veterans who had a mental health need and did not use VA mental health services, 33% did not know VA offered them, and 42% did not know how to apply.[36] Across every site the committee visited, veterans were confused about the difference between the Veterans Benefits Administration (VBA) and the Veterans Health Administration (VHA), two different organizations within VA that each operate their own programs. Many assumed, reasonably, that going through the disability compensation process, which involves undergoing medical exams through the VA disability claim process, had enrolled them in VA health care. It had not. They were talking to VBA, while VHA enrollment is a completely separate process with its own eligibility criteria. To a veteran, it’s all VA. Rational, but not correct.

[36] National Academies of Sciences, Engineering, and Medicine, *Evaluation of the Department of Veterans Affairs Mental Health Services* (Washington, DC: The National Academies Press, 2018).

Throughout this paper, we've talked about definitions diverging *across* agencies. Those are coordination problems, and coordination problems have the comfortable property of sounding solvable with better coordination. But VBA and VHA report to the same secretary, *inside the same agency*. They share a budget, a general counsel, an inspector general, a congressional oversight committee, and two-thirds of a name. There is no interagency boundary to blame for producing two incompatible meanings of "enrolled in VA."

Marina has long believed that the best possible experience a veteran can have of VA benefits is *no* experience. The day you separate from the military, we already know everything we need to know about what benefits you're eligible for. Money should simply appear in your bank account — *that* would be an experience that honors veterans for their sacrifice. There's no law of physics preventing this reality.

A front door improves the experience of people who already knew to show up; it doesn't help people who didn't.

Yes, sometimes we genuinely can't act until someone takes action. Social Security survivor benefits are one example: A divorced spouse can qualify on their ex's record if the marriage lasted 10 years. But SSA has no idea that marriage existed, so the people most likely to be missing this money are the ones who'd never think to ask about it. But most of the time, we already know everything we need, yet we wait for people to show up and ask anyway.

8. Combined, Dynamic Applications Alone Can't Work

The intuition behind a one-stop shop is that a person should only have to enter their information once. That sounds great until we try it in practice. Our colleagues Christopher Whitaker and Arabella Crawford do a great job documenting this challenge in *Standard Definitions*, which we build on in this paper.[37]

Here's a scenario:

> Dana was laid off at the end of February and now receives unemployment. Her husband, Marcus, earns $1,900 a month, paid every two weeks. Their daughter, Amaya, is 16 and makes about $380 a month at a part-time job. Their son, Theo, is 8, and presently unemployed. Dana's mother, Ruth, 68, lives with them and receives $1,240 a month in Social Security; she buys and prepares her own food. Marcus' brother sends $250 a month toward the rent. In January, Dana sold some stock.

Take the most common question in the safety net — *what is your income?* — and try to put it in a combined application. This quickly exposes the underlying challenge: According to the

[37] Arabella Crawford and Christopher Whitaker, *Standard Definitions: Reducing Administrative Burden by Ensuring Interoperability* (Washington, DC: Searchlight Institute, July 27, 2026).

government, the same household on the same day has at least six correct yet mutually inconsistent incomes:

Unemployment insurance wants Dana's wages only, from covered employment only, over four of the five most recently completed calendar quarters. Marcus' earnings are irrelevant. So are Amaya's, Ruth's, the rent money, and the stock sale. So is the unemployment Dana is currently receiving. The answer is a historical number about one person.

SNAP wants the current month, for everyone who buys and prepares food together. Ruth is out, because she cooks separately. Dana's unemployment is in. The program needs two numbers, gross and net, and the household must pass both tests. The net figure subtracts deductions for shelter and care costs, and those deduction rules vary by state.

Medicaid, on the MAGI pathway, wants modified adjusted gross income for the current month, across the tax filing unit. MAGI is adjusted gross income plus tax-exempt interest, nontaxable Social Security, and excluded foreign income. The tax unit here contains the same four people as the SNAP unit, which is a coincidence rather than a design: One is defined by who files together and the other by who cooks together. Ruth, if nobody claims her as a dependent, is her own unit, and her nontaxable Social Security counts toward her MAGI even though it would not appear in her adjusted gross income. The stock sale is in adjusted gross income, but the measurement window is the current month, so it lands in January's number and not in this one.

Section 8 wants annual income for all adult household members, projected forward 12 months from current circumstances. Ruth is back in, because she lives there. Amaya is out, because HUD disregards the earned income of children under 18. This is an anticipatory number about a year that has not happened.

LIHEAP wants household gross income, compared against a threshold each state chooses. Federal law lets states set eligibility anywhere from 110% of the federal poverty guideline up to the greater of 150% of the guideline or 60% of state median income. Some grantees count child support that others do not. Both the number and the line it is measured against depend on where the household lives.

The ACA marketplace wants MAGI, the same definition Medicaid uses, projected annually rather than measured monthly. Same rule, different window, different number.

Six programs, six numbers, and none of them is wrong. There are four things actually varying in the figure: who counts as the household (Dana alone, the four people who cook together, everyone under the roof, or Ruth by herself); which money counts (wages only, or wages plus unemployment, or the brother-in-law's gift, or a stock sale that counts for one program but not for the next); over what window (last month, the next 12 months, or the four quarters that ended before Dana lost her job); and whether the number is gross or net of deductions. Each program picked its answers for its own reasons, and none was ever asked to agree with the others.

Today, a combined application has only two options here.

It can **ask everything**, collecting all six programs' inputs so it can compute each answer correctly. This makes us look insane, and rightly so. It also ends up being longer than any single application, and longer than if you chained the single applications together, because the combined form has to establish who cooks and who files and who lives together, *before* it knows which programs the person may qualify for. The promise was one short form, but the end result may very well be a much longer experience than applying individually would have been.

Or, it can **ask once and make educated transformations** behind the scenes, collecting one plausible income figure and applying it to all six programs. This is the version that often gets built, because it demos best in a meeting. But it produces many wrong answers, and does so *invisibly*. The problem only surfaces later, when the applicant gets a denial, or a bill to pay back an overpayment (often via a scary red envelope) that they probably can’t afford.

There is a third option that is laid out in the rest of this paper: Don’t ask the person for income at all. If every program's version of "income" has been written down as a question a central verification system can answer (wages for whom, over what window, from what sources), and one authoritative government source holds the answers, then the combined application doesn't have to choose between asking everything and guessing. It can ask the source six questions and get six correct answers, without Dana having to do anything. The programs don't have to agree on a single definition, but they do have to agree on where to look. That's why reconciliation and an authoritative source come *before* the form, and not after it.

LESSONS FROM WHAT'S WORKED

The previous section may have depressed you. If we've tried and failed so many times, has anything we've done actually worked?

Turns out … yes!

Table: What's Actually Worked

Year	Effort	Rung	What happened
1966	Medicare Parts A and B auto-enrollment	6	Anyone already receiving Social Security at 65 is enrolled automatically. No application, no screener, no portal. SSA already knew their age, identity, and work history, so it did the thing.
1974	SSI to Medicaid	6	In about two-thirds of states, an SSA determination of SSI eligibility automatically confers Medicaid.
1981	Summer Food Service Program	6	The Omnibus Budget Reconciliation Act of 1981 (P.L. 97-35) redefined "area of poor economic conditions" to mean areas where 50% of children would be eligible for free and reduced-price meals, so Summer Food Service Program operators can serve free meals to all children.
2004	Direct certification for school meals	6	P.L. 108-265 required states to certify children in SNAP households for free meals with no household application. Nationwide by SY 2008-09. By SY 2018-19, USDA reported 98% of children in SNAP households were directly certified.
2006	Medicare Savings Program to Part D Low-Income Subsidy	6	Anyone enrolled in a Medicare Savings Program (or Medicaid, or SSI) is automatically deemed eligible for the full Part D Low-Income Subsidy. CMS does this from state enrollment files; the person never applies.
2006	Pension Protection Act of 2006	6	Law that gave clear legal backing and requirements for some employers (including the federal government!) to automatically enroll workers in retirement plans. While the military was excluded, evidence from the Office of Evaluation Sciences field tests showed service members enrolled readily when asked, so "opt-out" was included in the 2017 National Defense Authorization Act.
2010	Community eligibility for school lunch	6	Allows schools with more than a certain percentage of kids qualified for free lunch to provide all students free lunch.
2012	Ex parte Medicaid renewal	6	Federal rules require states to attempt renewal from data already held before sending anyone a form, with successful renewal rates ranging from 10% (TX) to 99% (RI).
2021	Advance Child Tax Credit	6	IRS delivered monthly payments automatically to most eligible families from tax data already on file.
2024	New Jersey ANCHOR property tax relief	6	New Jersey files the application on behalf of most eligible homeowners and renters, using residency, income, and age it already holds from tax records and prior filings, and mails a confirmation letter. About 1.5 million residents were auto-filed in 2024.
2024	Summer EBT / SUN Bucks	6	Children already enrolled in SNAP, TANF, or FDPIR are automatically eligible; states enroll them without an application.
2026	530A (Trump) Accounts	6	Dell Foundation made $250 deposits to every child (who has an account) born between 2016 and 2024 residing in a ZIP code where the median household income is below $150,000.

Table: Searchlight Institute • Source: See Appendix E for citations.

There's a throughline that ties all these projects together: These are all Rung 6 examples.

Not one of these projects was called "integrated benefits delivery." None had a portal, a launch event, a systems integrator, or a 21st Century "NextGen" modernization initiative. In fact, they were not really technology projects at all. *They were policy decisions.*

Colorado wasted $100 million on CBMS and Oregon $240 million on Cover Oregon, both trying to reach Rung 5 by building a technology system. In contrast, automatically enrolling everyone already on Social Security into Medicare at 65 has worked for decades at a (technology) cost of virtually nothing.

Are these perfect models? No. Several states declined to participate in Summer EBT. Ex parte renewal rates vary enormously between states. The expanded, monthly version of the Child Tax Credit lapsed after a single year. But they are still proof that this works when we choose to do it.

What every example has in common is that in each case, one part of government held all the facts needed to make a decision, knew the rules to apply, and had permission to act on both. SSA knows your age and your work history. The IRS knew the family's dependents and ACH details.

That's the entire trick. Nothing about these wins required a new website. It's also a usable test. Before anyone builds anything, ask four questions: Does some part of government already hold the facts? Does it know the rule? Does it have the authority to act? And does somebody own the outcome if it acts wrong? Where all four are yes, the rest is plumbing. Where any one is no, *that's* actually the work, and no amount of plumbing substitutes for it. Every failure earlier in this paper was a "no" on at least one of those questions.

The question is not whether the top rung is reachable. We have been able to reach it since at least 1966. The question is what it would take for the top of the ladder to be the norm, rather than a rare occurrence — to have the facts, the rules, and the permission in place across programs, rather than hoarded inside them.

Every example in the table got there one of two ways. Either the program's criteria were already facts the government held (your age, your Social Security record, your SNAP enrollment, your Medicare Savings Program enrollment), or somebody rewrote or adjusted the criteria *until they were*. The Summer Food Service Program stopped asking about each child's household and started asking whether half the children in the area qualify; community eligibility did the same for the whole school for students' access to free breakfast and lunch during the academic year. That's the move we want to make. Keep the criteria an authoritative source can answer, and where a criterion can't be answered that way, move it to a population-level proxy, rewrite it as something a source can in fact-check, accept the person's word for it, or consider whether the program needs that criterion at all.

HOW TO CLIMB THE LADDER

In order to get to the top rung, the facts, the rules, and the permission have to align across programs instead of inside each one. This future has to be governed into existence, and we can take steps to get there even now.

Step 1: Consistently Document Eligibility Criteria

Starting *today*, we can write down the eligibility criteria for every benefit administered by a core group of federal providers (at least SSA, USDA, HHS, HUD, DOL, IRS, and VA, and probably SBA, FEMA, and USCIS), in one template, including the source of each requirement and where states, Tribes, and/or counties layer more criteria on top. Most of this work has already been done, to a high level of fidelity, by people outside government.[38]

This is not a recommendation to issue an executive order directing agencies to submit an inventory. Instead, we should compile the inventory in advance, and confirm its accuracy to the best of our ability. On Day 1, the president directs agencies to confirm the accuracy of the compilations by a short deadline (weeks). If they don't respond, they've agreed to whatever we wrote (and if an agency would rather not respond at all, Congress can ask for the same information under subpoena). Congress sees the resulting compilation, as does the public.

This cannot be a voluntary, opt-in effort.

Notice what's *not* in this step: giving up a criterion, changing a threshold, accepting somebody else's definition, handing over a form, or losing a dollar. This step is specifically about agencies writing what they are ostensibly already doing and, one would think, have already written down.

Donella Meadows describes the way to handle a massively complex system: Break it into subsystems, each with a manageable set of inputs and outputs, so that the work inside any one of them can be done without holding the whole in your head.[39] "The whole of federal and state benefits eligibility" is overwhelming. But "confirm this table for this program by March" is ordinary work.

We already run this model at an enormous scale, once a year for taxes. The IRS publishes the forms, instructions, e-file schemas, and business rules that every tax product has to submit against. TurboTax, H&R Block, Free File, and your accountant's software all read and conform to

---

[38] Sonal Ambegaokar, Zoë Neuberger, and Dorothy Rosenbaum, *Opportunities to Streamline Enrollment Across Public Benefit Programs* (Washington, DC: CBPP, November 2017); Ariel Kennan, Lisa Singh, Bianca Dammholz, Jason Yi, and Keya Sengupta, *Exploring Rules Communication: Moving Beyond Static Documents to Standardized Code for U.S. Public Benefits Programs* (Washington, DC: Beeck Center for Social Impact + Innovation and Massive Data Institute, updated September 4, 2025).

[39] Donella H. Meadows, *Thinking in Systems: A Primer*, ed. Diana Wright (White River Junction, VT: Chelsea Green, 2008). Meadows is relaying Herbert A. Simon, "The Architecture of Complexity," Proceedings of the American Philosophical Society 106, no. 6 (1962): 467–82.

the same published rules. The IRS publishes what a valid return looks like and lets everyone else build interfaces based on those rules.[40]

Specify the verification source of every criterion

The template for eligibility criteria across programs has to include all information requested to determine eligibility for and deliver each benefit, whether or not the government ultimately verifies it.

Some facts are currently only produced through the application process, like a current rental agreement, attestation from a landlord, or list of people with whom an applicant purchases and prepares food. Some facts may come from a third-party vendor today. Each eligibility requirement needs a precise specification: what specific fact or calculation is required, from whom, for what time period, how it's currently verified, and, just as importantly, how much verification the rule actually requires (known as level of assurance).

Much of what applicants are asked for isn't actually in any rule. Federal SNAP regulations let a household self-attest as to who buys and cooks food together, unless something looks questionable.[41] Yet many states and caseworkers demand a landlord's letter or a neighbor's signature for every applicant, anyway. That's oververification — requiring more paperwork than the federal regulations demand — and it's the cause of a large share of procedural denials: The person was in fact eligible, but the case was denied because a landlord letter didn't arrive (or, more likely, the landlord is a private equity firm that has no idea with whom you cook food). Documenting the required standard next to the actual practice makes the gaps between the rule and the practice visible, which is the first step to closing them.

Underneath all of this is something about eligibility determination that we rarely think about: A determination is a legal finding. Every yes or no has to potentially survive a fair hearing, a quality control review, an inspector general, and even a judge. That's ultimately why the caseworker asks for a landlord letter: A letter in the file is proof they can point to if the case gets pulled, whereas "the applicant told me" feels less defensible even when regulation allows it. Oververification is defensive documentation.

For each criterion it has to say what evidence is legally sufficient to defend the determination later: the regulation that makes self-attestation acceptable or the data match that counts as verification; how long that evidence stays valid; and the standard a hearing officer will apply. Once that's written down, we can start designing a world where a caseworker has permission to stop collecting extra steps (and perhaps the system outright prevents them from doing so in the future).

---

[40] The part of that story that went badly isn't a reason to stop publishing the rules.

[41] New York City Human Resources Administration, Family Independence Administration, *Supplemental Nutrition Assistance Program (SNAP) Documentation Guide*, Form W-129G (New York, NY: New York City Human Resources Administration, August 21, 2012).

Payment error rate penalties are making this worse. A state that gets fined for paying the wrong amount, but not for denying benefits to an entitled recipient, will rationally pile on paper in defense. For facts only a household can supply, the rule should say self-attestation is the default, and explicitly detail what triggers anything more. Burdens that can't be automated have been intentionally added to certain benefits, in order to shrink the rolls, and then more measures to try to un-shrink the rolls get added by the other side, ad nauseam.

Let's have the argument about whether and to whom we should provide SNAP out in public, rather than trying to "win" via a pile of thousands of landlord letters. This template is where those debates can originate.

Document the discretion, not just the rules

A natural objection is that documenting eligibility criteria is much harder than it sounds. An early reader of this paper put it bluntly: Agencies know, and will want to protect, that ambiguity is power. A lot of eligibility language is vague on purpose, and that vagueness is where discretion lives — the caseworker's judgment, the regional adaptation, and the ability to deal with a family that does not fit neatly into the original table of eligibility criteria.

Asking an agency to squeeze all this into a rigid table *is* asking it to give up something: the flexibility it legitimately uses every single day. If agencies won't voluntarily give up control, and they'll maliciously comply if ordered to, then how do we generate the accurate baseline of facts we need if we are to build a truly integrated benefits system?

Ashby's Law recognizes that a rigid criteria table is *also* a regulator. A table has less variety than the entire population of American households, so it can't absorb all of them, and the variety it can't absorb comes back out as denials, appeals, and lawsuits.

Caseworker discretion is not an obstacle to our goal. Rather, it *is* the variety absorber, and right now basically the only one we have. Vagueness in regulations does real work in (hopefully) helping households that original statutes did not address. Strip it out and we have yet another rules engine with insufficient variety — one destined to fail.

This first step is not about eliminating ambiguity, but rather, it is a request to document where the discretion lives and what bounds it.

This can start as a column in the template. For every criterion: Is this rule-bound or judgment-bound? If judgment-bound, who holds the authority, what are the boundaries, and what gets recorded when judgment is exercised? An agency can fill this out while keeping every ounce of the judgment it currently exercises; what it gives up is the *invisibility* of that judgment.

Right now a lot of discretion is unprotected. For example, a SNAP caseworker is supposed to decide whether someone is "obviously unfit for work," but many don't believe that's in their job

description, and are afraid of what could happen to them personally if their call is second-guessed later. As a result, they don't make the judgment call at all, and the person who needed the exemption doesn't get it. Documenting who holds the authority is also documenting that they're *allowed* to use it.

This exposes another kind of problem, where a vague criterion is quietly carrying an outsized amount of weight. "Medically frail" under Medicaid work requirements was supposed to catch the edge cases but ended up catching everybody. **Rules should be clear and automatable for most people, with discretion available for the households that need it.**

The template also has to be versioned. Every layer changes, and more often than we think: Congress amends a statute, an agency issues guidance, a state legislature changes a threshold, a county supervisor issues a memo, we find and fix an error, and so on. Rules kick in on one date and sunset on another, while old rules have to be retained for audits, appeals, and retroactive claims. This catalog is meant to drive action, not become a museum exhibit.

The output of all this is a complete, current, public list of what the rules actually are, how they're verified, and where they're black and white versus gray.

For Step 1 to work, the president must also get the Paperwork Reduction Act (PRA) out of the way. As currently interpreted, the PRA requires OMB clearance before a federal agency can ask the state and Tribal agencies *that administer its own programs* to give it the eligibility criteria they apply. A reconciliation effort that has to route every question to states through a multimonth clearance will never succeed.

While this is ideally a legislative fix to remove the information collection requirements from the PRA altogether, a Day 1 action could be an EO declaring use of the existing PRA "pilot" authority. The pilot is: What happens if we stop applying the information collection requirements of PRA? Stay tuned to find out.

Step 2: Establish and Enforce One Authoritative Source for Each Criterion

Integrated benefits requires one authoritative source for a given fact: one that's readily accepted by all other benefit lines after it's confirmed once. Once this source is established for a given criterion, government programs are not allowed to ask for or verify this fact independently. "One source" doesn't mean one database. For a fact like wages, the data will always live in several places with different coverage and timeliness. It means one service that owns the job of stitching those together, stands behind the answer, and is the only place anyone is allowed to ask.

Start with income

This has a clear solution and a huge impact, following the (eventual) success of the United Kingdom's Universal Credit program, and similar programs in many other modern countries.

The raw materials for a national income verification service exist and are already funded, but they are too slow and siloed. State unemployment agencies already collect wage data from employers, but on a *quarterly* basis. They rarely share this data with other agencies, even in the same state. The Institute for Responsive Government has one blueprint for how America could fix this.[42]

Building a national income verification system could involve the following steps:

1. Require state unemployment agencies to collect wage data on a per-paycheck basis as a condition of receiving their share of federal funds. DOL recently issued a letter to state UI directors with a footnote encouraging investment in more frequent payroll data collection.[43]
2. Require state unemployment agencies to make this income data available to other government agencies as a condition of receiving their share of federal funds.
    a. This could be designed as an API that accepts a few variables, like "over what time period." It could protect privacy by simply returning a "yes" or "no" response to whether the calculated amount is over or under the requested number.
    b. Example: "Did Dana make more than $200 in the last 14 days?"
        i. Response: "No" (instead of responding with Dana's payroll records or her specific income in the last 14 days).
3. Roll this data up to the IRS for a national income verification system that has data as recent as everyone's last paycheck. The IRS ends up with this information through taxes, so it's not collecting or storing anything new. (A side effect: The government could see labor market shifts as they happen, paycheck by paycheck, instead of months later via a survey.)

Work will be needed to include other forms of income beyond employer-reported wages, and to make the wage records themselves consistent, since what counts as a reportable wage differs from one state's UI system to the next. But this is tractable, as evidenced by the third-party vendors who are selling income back to the government already.

One thing that makes criteria like income harder than something like citizenship is that a means test asks you to prove you *don't* have something. A source that confirms your Walmart wages is only useful if it can also confirm you don't have Uber wages. This is why the source has to be complete, not just accurate, and why the IRS is likely the right long-term home: It is the one place with both the coverage and the audit function to verify completeness.

---

[42] Hannah Garden-Monheit, Dan Meuse, Gnora Mahs, Marisa Bremer, and Rebecca Piazza, *Modern Income Verification: High-Tech Integrity, Low-Cost Delivery* (Washington, DC: Institute for Responsive Government, May 5, 2026).
[43] Luke Timmons, Letter to State UI Directors Sharing Model National Unemployment Insurance Standards (Washington, DC: U.S. Department of Labor, Employment and Training Administration, July 6, 2026).

And then the part that makes this all work: **Once this service exists, agencies may not request income, or verify income, any other way outside of the official source.** A shared service that agencies are merely *welcome* to use is a shared service that agencies won't use.

The "income" fields on the forms go away. The pay stub uploads go away. The task of verifying someone's income is permanently checked off. The third-party income verification contracts expire. Tax filing is potentially a dramatically simpler process. This could even become a service for verifying income for private sector purposes, such as mortgages and car loans.

This service would have to be resourced and held to high standards of accuracy, availability, and ease of use, but we have better odds of doing this well in one place, than in many.

Identity: accredit it, don't build it

The government already issues the identity documents that matter: driver's licenses, passports, and Social Security numbers. What it doesn't need to build is the system that turns those into a digital credential. The government can operate its own digital identity system, or it can accredit providers against a set standard. The Kantara Initiative already performs this third-party conformity assessment, having certified ID.me, CLEAR, and Login.gov. The government should accept a verification from any certified provider at the assurance level the transaction warrants.

That means an end user can connect their credential to whatever user interface they choose to use to access benefits. That credential is digital and portable: proved once, held by the provider, and presented to any agency that accepts the standard, the way a passport works at any border. Ideally, I can use my credential to carry certain verified facts with me (like a backpack).

This prompts a series of hard questions, like: What ties a verified identity to the right income record, or the right death record, across benefit lines? This is a real problem with multiple potential solutions, and a big enough one that it will be the subject of a separate upcoming paper.

A catalog of authoritative sources

As this effort to standardize criteria succeeds, a central list should be published that details which facts exist and which authoritative service can confirm each one.

For every fact a program might need to verify — date of death, citizenship or immigration status, wages, incarceration status, veteran status, Medicare enrollment — the catalog records four things: what the fact is, who the authoritative source is for it, the mechanism for asking, and whether the answer is permanent or changes over time. (Note: The catalog holds no data about any person.)

That fourth item — whether the answer is permanent or changes over time — matters for reasons described earlier; date of birth never changes, but wages are true only as of a specific date.

The authoritative sources also need to account for correction. A fact that "never changes" can still be incorrect and need updating. For example, about 10,000 living people a year get added to the Death Master File by mistake. Every source has to offer a way for a person with authority to correct the record and push that correction to everyone who relied on it.

Many of these services already exist

U.S. Citizenship and Immigration Services (USCIS) runs SAVE for immigration status verification. SSA maintains death data. The Office of Child Support Services runs the National Directory of New Hires (NDNH), which already holds quarterly wage and new-hire records. The VA can verify military service history.[44]

Most data sources are not federal. Wage records sit with 53 state UI agencies. Residency and address usually sit with a DMV. State income tax lives with the state revenue department, and often, so does the best current mailing address. "Anoint the authoritative source" may refer to a state agency, on behalf of every federal and state program that needs the fact (or, where possible, one federal source queries all the approved state sources).

But existing isn't the same as working. States are already required by law to check the NDNH when verifying income for SNAP, and the ones we've talked to say it rarely helps because the data doesn't answer SNAP's question: NDNH holds quarterly wage totals and new-hire reports collected for child support enforcement, while SNAP needs this month's income, by employer and by pay date. When SNAP tried to reuse the Centers for Medicare and Medicaid Services (CMS) data services hub, CMS wasn't funded to treat SNAP as a customer, so it didn't. Nobody at either agency did anything wrong; they were not set up or resourced to cooperate.

So anointing a source must come with at least three things:

1. Direct funding from Congress to operate the source at a high level of fidelity for everyone who depends on it, instead of the cost-recovery fights that eat a year and produce a pointless MOU.
2. A mandate that says serving other agencies is *the* job, not a favor. It needs to be resourced to support other agencies, with formal accountability for their needs.
3. Accountability for whether the data actually answers the questions the dependent agencies are asking. "We have wage data" and "we have the wage data SNAP needs, when SNAP needs it" are different. It's hard to convince the DOL to spend its own money to save SNAP's and Medicaid's; somebody with a cross-agency budget has to be able to make that trade.

[44] U.S. Department of Veterans Affairs, *VA API Platform: Verification* (Washington, DC: U.S. Department of Veterans Affairs).

That's what we need: **the anointment of each as the authoritative source, the funding to be one, and the firm edict that you don't get to independently ask or verify on your own, or your program will be found out of compliance.**

For single, government-run authoritative sources to work, we'll need to modify and reconcile some existing criteria to be able to use them.

Step 3: Reconcile Criteria

"The different programs are defining and measuring details like income completely differently and they first have to be reconciled before a technical fix can be implemented." - Chris Whitaker and Arabella Crawford

With the criteria written down, conflicts become clearer. Some differences are deliberate policy, whether for good or nefarious reasons. But many are accidents, because programs develop their own criteria and definitions in silos. In any case, we need to do the work to reconcile them, both to simplify the underlying eligibility layer and to enable the use of more authoritative sources.

Reconciliation is at least five different things

"Reconciliation" isn't one activity; there are at least five distinct flavors. The first three are about making a criterion answerable at all. The last two are about whether programs should want the same answer.

1. Can the authoritative source even answer the question? A criterion has to be expressible as something you can ask a system. "Made less than $X in the last 30 days" is an objective query. "Made less than $Y in the last 365 days" is a different question, but still an objective query. A two-page narrative in a policy manual about how income is assessed, with exceptions influenced by which caseworker happens to pick up the file, is *not* a question an authoritative source (meaning one system of record such as SSA for date of death) will ever be able to answer automatically.

This is the minimum bar. Rewriting a criterion so that a system can answer it may well require statutory changes, but these should be budget-neutral, because it's "only" a restatement of (in theory) current eligibility criteria.

Some criteria will fail this test no matter how they're rewritten, and that's useful information. A criterion that can only ever be verified by a document the applicant has to go find costs the applicant, the caseworker, and the system as a whole time, and its advocates should have to justify that cost. The options: Move it up to a population-level test, replace it with a proxy a source can check, accept self-attestation, or drop it. Caseworker discretion can stay for the households that need it, which is a much shorter list than the one we have today.

The goal of this reconciliation step is that everyone can query the same authoritative income service for their authoritative answer. The goal is *not* that everyone's income eligibility becomes 140% of $X over the last 365 days.

2. What counts? Does a Connecticut Baby Bond count as income? (Hopefully not, since they were explicitly designed not to.[45]) A neighbor who bought you a pizza? Your parents' assets, if you're 19 and living at home? VA disability compensation counts toward eligibility in some safety net programs and not others; RAND has shown that difference alone drives participation rates.[46] Every program has a list of inclusions and exclusions, which adds significant complexity.

Another detail: proxies. In some places one criterion is already standing in for another, like income used as a proxy for hours worked, or enrollment in one program used as a proxy for eligibility in another. Existing accepted proxies are a form of reconciliation that somebody already did. What counts as a proxy now? Are there available, unused proxies a benefit could leverage? Do we care if it leads to gambling?[47] Where else might an existing proxy be used?

3. Over what period, and for how long? Is income evaluated for the last 30 days, last month, 365 days, or four of the last five completed quarters? Unemployed for how long before you qualify? Calendar days or business days (and if business days, what holidays count)? And on the other side: How long does the benefit last, and how often do you have to recertify a given fact? Could we reconcile that timeline to align with other benefit lines or a more widely used timeline (e.g., a calendar year)?

4. Should the thresholds actually be the same? This is where "reconciliation" turns into "alignment." Maybe SNAP and LIHEAP should have the same income limit, or maybe they shouldn't! This is a real decision with real enrollment consequences, and we can make strategic decisions about where to focus alignment attention once everything is quantified in the above steps.

5. The impact of the benefit itself. Once someone receives a benefit, what happens to it on the way back into the system? Is it taxable income? Does receiving the benefit count as income for the next program? For example, foster care maintenance payments are not taxable, but if that foster home takes in another foster child for the weekend as "respite," *that* income *is* taxable. This is very confusing, and creates a structural disincentive for anyone to provide respite.

Once you reconcile a criterion (e.g., what counts as income, and over what period), that criterion should be readily *understood* across benefit programs, and (assuming it's validated and timely)

---

[45] Connecticut Office of the State Treasurer, *Sample Participant Letter* (Hartford, CT: Connecticut Office of the State Treasurer).

[46] Philip Armour, Catria Gadwah-Meaden, and LeeAnn Stevens, *Veterans' Participation in Public Programs: Comprehensive Cross-Survey Comparisons with Nonveterans and the Role of VA Disability Compensation*, RR-A4405-5 (Santa Monica, CA: RAND Corporation, July 28, 2026).

[47] Kevin Corinth, *The Medicaid Work Requirement's Gambling Loophole* (Washington, DC: American Enterprise Institute, September 15, 2026).

can be reused across a greater number of programs without manipulation or reassessment. For any given criterion, the goal of reconciliation is that it covers the whole of government. In the real world, reconciliation may happen across *part* of government, with a battle plan drawn out for the remainder.

Reconciliation will need to be driven by a central team. To be successful, it must be a collaborative effort to understand a program's goals and needs, and determine the best path forward for all involved. The team will need sticks; for example, once the federal government builds an income verification source, over the five years after it launches, the block grants given to states to pay for income verification should phase out to zero. That is: Once there's one income source every program can query, the money states get today to buy or build their own income checks goes away. It doesn't necessarily mean there's only one feed behind the source; state UI wage records and IRS tax data can both feed it. It does mean there's one place to ask.

A reconciliation effort can provide carrots, too. An agency named as the authoritative source for a fact gets funded to operate that source permanently, and to answer everyone else's queries well (with a documented, required level of service). Today, being the system of record mostly means doing extra work for other agencies and chasing them for cost recovery. This would flip that incentive.

This step somewhat overlaps with Step 2: As you reconcile criteria, some decisions may be influenced by where and how you can automatically and centrally verify that fact.

One column in the template should include an "implications of change" field, in order to capture where even a slight adaptation to the definition could have cascading impacts on how people are served, like "donut holes" that may be created, or benefits cliffs that create perverse incentives or can materially harm families. (This is where rules as code can come in handy, by automatically identifying conflicts as rules are updated.)

This is a transformative effort, and not everyone will benefit. Every reconciled definition moves a line, and some recipients will be left on the other side. If we want, we can track and avoid disenfranchisement along the way, because the definition and the threshold are separate dials. If a program changes what it counts as income or over what period, it can reset its threshold so roughly the same population still qualifies. When a program chooses to serve fewer people, let's make that decision on purpose, out in public.

Conformance testing (meaning that a system's answers match a published set of test scenarios, described in Step 4) forces the same conversation from the other direction: Some state systems will turn out to have been more generous than the statute, and correcting the code unenrolls innocent people. A hold-harmless period, where anyone enrolled under the old rule keeps their benefit through their next recertification, would buy time for those households, and for the state to decide whether it wants to change the actual rule.

Identify the first tranche of reconciliation steps

What criteria cross the most programs? Where are criteria set in legislation (harder to fix) or regulation (easier) or policy (easier still)? Where have agencies indicated they're willing to make changes? It's entirely possible to come up with a first pass of changes that can be made in the first six months, *and make them*.

Conversations with leaders who have been working at this problem for years revealed there may be very little reconciliation juice left at the agency level in high-visibility programs like SNAP and Medicaid. Multiple working groups in the last administration tried to align those two programs' rules and found that what remains requires Congress. But it's still worth looking at programs like LIHEAP, Head Start, WIC, and the dozens of smaller programs where fewer people have paid attention, to see if opportunities for streamlining exist within the executive branch. We can start to identify these opportunities now.

Marina saw a willingness to reconcile firsthand when she helped write new standards for licensing foster homes.[48] In foster care, each state defines "relative" many different times, including for emergency foster care placement, kin-specific foster home licensing, notifying when a child enters (or is at risk of entering) care, child-only TANF, calculation of the numerator of the penetration rate for title IV-E federal financial participation, ICPC, guardianship and/or subsidized guardianship assistance, adoption and/or subsidized adoption assistance, and kinship navigator supports. But nobody *intended* or *wanted* to have multiple, conflicting definitions. Once everyone saw all the conflicting definitions next to one another, and we derived one recommended definition that encompassed them all, reconciling definitions across and within states was not a fight.

All criteria are similarly mummified in forms, website applications, and internal benefits processing systems — so the objection that those changes will be "hard" is heard and summarily rejected. Those things all change, all the time, and this will be one more change order.

Once the voluntary reconciliations have been identified and resolutions are underway, we move to the harder part of the work: forced reconciliation. This should ideally be prioritized by impact, with the criteria covering the greatest number of programs and recipients prioritized first.

In theory, lawmakers should start by reconciling "income." This effort will mostly be legislative in nature (though there are some definitions within the control of the administrative state, perhaps even more if the president is bold). Take all disparate definitions and work with the impacted benefit lines to **start to align them**. This should have a rapid execution window, and much of the planning work can be (and is currently being) done ahead of time.

[48] Marina Nitze, *How to Peel a Banana* (Eating Policy, September 21, 2026).

Then repeat this effort with other shared, harder-to-reconcile concepts like “household” and so on.

Dealing with intentional divergence

Reconciliation can solve the problem of programs using one word for several slightly different things. It cannot “solve” for programs legitimately asking different questions, but using the same term. Knowing the difference keeps the scope of early reconciliation reasonable, because it doesn’t ask an agency to give up control over the criteria it most cares about defending: the ones that *are* its policy. Later, we can take up the question of whether VA disability compensation, Social Security disability, workers’ compensation, and unemployment insurance *should* all have different criteria for “disability.” That's a different fight, and we can pick it later from a better position.

The hardest part of harmonizing criteria is where legislators or policymakers deliberately wrote criteria without setting clear, objective standards. Nobody added a footnote noting which gaps in definitions were negotiated and which were a well-intentioned workaround from a supervisor. Each gap gets equally defended as sacred, while the people trying to apply them feel like they’re reading tea leaves. The template's "source of requirement" column can help: A definition traced to a floor amendment is a different conversation from one traced to a 1987 policy memo.

Vagueness isn't only a tool of caseworkers and agencies. It's also a tool for legislating. Congress (and states) write imprecise eligibility language on purpose all the time, and sometimes the imprecision is the only reason a law gets passed — a definition intentionally left vague so two members who disagreed on the details could both vote yes.[49] When that's what happened, the gap between two programs' definitions isn't an accident of siloed drafting at all. Somebody did it on purpose.

Some divergence is defensive. Take the SNAP “household.” Your eligibility depends on who you purchase and prepare food with. No other benefit line uses this definition of “household,” and for a lot of households it’s a genuinely difficult question to answer.[50] But the “household” definition is weird partly because it was written so as not to implicate the household definition used for SSI; if SNAP's household matched SSI's, a change in one could cascade into the other and knock people off both programs. This is why the "implications of change" column exists, and why cliff-mapping is so important.

---

[49] That bargain got shakier in 2024, when the Supreme Court’s *Loper Bright* decision ended courts’ deference to agency readings of ambiguous statutes. A definition Congress left loose for the agency to fill in is now filled in by a judge, making one more reason for Congress to write the definition it actually intends.

[50] The answer can depend on who's staying where and whose shift ended when. Nobody from the government can truly verify it without watching live video surveillance of your kitchen; in practice, it means a caseworker interrogates somebody about their domestic arrangements instead of building rapport with them.

And some divergence is an artifact of Congressional Budget Office (CBO) scoring. Any legislative change that makes a program easier to administer, or aligns it with another, will be scored by CBO as costly, because CBO scores expanded access as cost. The counterfactual is also true – making a program harder to administer can be scored by CBO as generating savings. In the case of the 2025 SNAP changes, CBO's savings estimates incorporated some likelihood that states may find administration so onerous they'd drop out of offering SNAP altogether.[51] This is a reason many advocates have been unwilling to support efforts to legislatively align programs: They know aligning criteria would be more efficient, but assume Congress would use the most restrictive and exclusionary definition to achieve a "good" budget score.

Some of what lowers the cost of benefit delivery today is the explicit assumption that some people who qualify will never receive the benefit. Roughly 1 in 5 people eligible for SNAP don't receive it; for WIC it's closer to half. Every budget today depends on that underenrollment. **Imagine if CBO learned to score divergence itself as expensive, since every new definition that departs from a standard test costs something to implement. Right now that very real cost is scored as zero.**

Our proposal saves money spent administering the negative-value gaps: reverification, churn, procedural denials, and the appeals that follow them. For the first time, a program would cost what the law says it costs. If Congress wants a smaller beneficiary population, then it can legislate one. It shouldn't get one by manipulating programs with intentional, unnecessary complexity.

It's not realistic to think *every* definition and criteria can be aligned. But at least we can document the discrepancies, quantify the harms, address them as best we can, and keep them on the list for a future where committee members change and agendas shift such that new reconciliations may be possible.

Reconciliations to make determinations (more) objective

Thus far, our proposal has been focused on improving the applicant's side of the experience. The other half of effectively integrated benefits is about the humans still making determinations when the computers can't. We'd only be doing half the job if we didn't address their experience, too.

Eligibility determination should be made as objective and as easy on the caseworker as possible.

Another part of reconciliation, which agency teams can largely take on directly if given good examples and support, is the user interface of benefits eligibility systems for the humans evaluating the applications. Information should be transformed and presented in the specific order and form the processing worker needs to make the determination in front of them, as

[51] Congressional Budget Office, *Estimated Effects of Public Law 119-21 on Participation and Benefits Under the Supplemental Nutrition Assistance Program* (Washington, DC: Congressional Budget Office, August 11, 2025).

opposed to simply capturing a PDF of the applicant's submission. We did a great deal of this kind of improvement work at VA and in foster care, and it is enormously effective.[52]

The presentation has to distinguish rule-bound criteria from judgment-bound ones. Take income, for example. Today, a caseworker may correctly object that they can't use income from any authoritative source, because it might be incomplete: Equifax shows the Walmart job but not the Uber driving. That's verifying an *omission*, which is the hardest kind of task, and why "automated income verification" today is so hard. But once the income source is *complete* rather than merely *accurate*, which is the whole point of having one authoritative source per criterion, the worker doesn't need to see the income at all — and probably *shouldn't*, lest they be tempted to reverify it or act on it.

One caution, because this is close to something we have spent the whole paper warning against. Building better software for caseworkers who apply incoherent rules makes the current system tolerable and extends its life. Building that software as the delivery mechanism *for reconciled, published, tested criteria* is the positive version. It is the same interface work. The difference is entirely in what sits underneath it, which is the argument of this paper in miniature.

Step 4: Set Up an Iterative Validation Framework

Every benefit needs to maintain a set of test scenarios that any system implementing its rules must pass. This would quickly surface any deltas between benefits processing system rules and the rule of law, and could even serve as a basis for performance-based contracting.

In practice, this is a list of realistic hypothetical scenarios. For SNAP, a test might confirm the eligibility and amount for different households, like a family of four with $2,900 a month in earnings and $1,400 in rent; a single adult paid every two weeks who got three paychecks in one month; a grandmother who cooks separately from the family she lives with; and a household where a teenager's part-time wages don't count. Hundreds of these, each with the correct answer listed, published by the agency that owns the program. SNAP currently has a test framework like this for its quality control process.[53]

Passing these tests has to be a system (or federal funding) requirement across all programs, and the tests need to run regularly. If a system doesn't pass the test, either the system is wrong or the test scenario is, and somebody with authority needs to decide and fix it.

Anything the criteria (e.g., the authoritative data source on income) do not clearly resolve should escalate to a human with authority to exercise judgment and decide. No criteria set will ever

[52] Quaylin Dang, *FormFest 2024 Recap: Scaling Government Assistance* (Washington, DC: Beeck Center for Social Impact + Innovation and Digital Service Network, March 24, 2025).
[53] Ariel Kennan, Elizabeth Bynum Sorrell, Jason Goodman, and Rachel Meade Smith, *Implementing Benefits Eligibility + Enrollment Systems: Insights on State Approaches and Processes* (Washington, DC: Beeck Center for Social Impact + Innovation, Digital Benefits Network, and Digital Government Network, February 4, 2026).

cover every last real case; the solution is not to pretend otherwise, but to consider and address edge cases explicitly. Situations that pass to humans should get evaluated as a new test scenario, if it's possible for the same situation to be handled automatically next time.[54] But some cases really are one-of-a-kind or require human judgment, and the right answer is that a human should *keep* deciding them. Those get marked as discretionary. What matters is that somebody *decides* a case is discretionary. If the same "unique" situation keeps showing up, we've lost the plot.

This design does not take caseworkers out of the loop. Instead, it allows them to document uncommon cases and, when possible, translate them into test scenarios that every implementing system has to pass from then on out.

Navigators and community organizations perform a version of this function informally today. When a rule is misapplied, usually a caseworker or a navigator notices. In cybernetic terms they are an "algedonic channel" — a fast, informal pain signal that reaches someone who can act, bypassing the formal reporting hierarchy that would normally take two years to surface the same issue. Formalizing that channel inside the system is how you keep improving and serving people as the volume of manual assistance goes down.

This kind of automated conformance testing is also a more accurate version of what the "payment error rate" efforts are trying to do. The error rate samples cases after the fact, at enormous cost, and produces an estimate so noisy that a state at 8.5% cannot be distinguished from one at 10.9%. Errors that erroneously gave people extra benefits are held equivalent to errors that denied hungry toddlers food.

Anyone who cares about program integrity should want conformance testing even more than those who care about ensuring access. Someone deliberately misrepresenting their circumstances is swimming in a very large pool of honest people whose answers *also* do not match, and every enforcement resource spent sorting out the honest mismatches is a resource not spent on investigating the actual fraud. Every definition we align, and every source made authoritative, raises the signal-to-noise ratio. Fewer honest people flagged, with the remaining discrepancies actually meaning something. That's a better deal for investigators than yet another data match service, and is a much more powerful defense against organized fraud.

It also ends the weakest-link game. Today someone lying about their income may only have to get past the program with the sloppiest verification; with one authoritative government source, there's one place to get past, and one place for the government to notice.

[54] The test scenarios will need to cover the collective donut holes and benefit cliffs, not just the individual program rules. A system can implement SNAP correctly and Medicaid correctly, and still push a household off a cliff neither program can see individually.

ALL FOUR STEPS, 53 TIMES

Everything above is written as though the federal government primarily administers these programs. But it mostly doesn't. SNAP, Medicaid, TANF, UI, LIHEAP, WIC, and child care are federally funded and state-run, and some programs like SNAP and foster care are also county-run. Tribes administer their own WIC, TANF, FDPIR, and Head Start programs.[55] The federal government controls some rules (and generally, the purse strings), but it's turtles all the way down. Let's think about how to scale this effort across all levels of government.

Documenting the layer cake

All of these programs have a state layer to contend with, and as long as they're federally funded or regulated, the state rules are in scope.

The template needs a jurisdiction dimension: the federal requirements documented once, and then each state documenting its options, additions, and policy manual interpretations on top. This is not impossible; FNA already publishes the SNAP State Options Report, which catalogs 28 choices across 53 jurisdictions.[56] Documentation largely exists today across many programs.[57] What we need now is the equivalent for every state-administered program, at the level of individual criteria rather than broad options.

The steps above are extra important for consistently capturing and validating consistency across so many implementations.

Reconciling in two directions

State-level reconciliation runs two ways: vertically, comparing the state's additions and choices against the federal criteria, and horizontally, comparing a state's own choices across the programs it administers, like Minnesota's SNAP options and criteria against Minnesota's Medicaid options and criteria. Because state options for different programs are usually chosen by different offices, years apart, with nobody looking at them together, there will be some juice to squeeze here.

Validating 53 implementations

[55] Tribal governments are not states, and nothing in this paper should be read as treating them as one more administrative layer. WIC, Summer EBT (SUN Bucks), FDPIR, Head Start, TANF, and other programs have Tribal options or components, and aligning any of them with the recommendations here requires government-to-government consultation under the federal trust responsibility. The documentation and validation steps may be the easiest place to start, since they describe what a program does rather than changing it, but that is for Tribes to decide.

[56] U.S. Department of Agriculture, Food and Nutrition Administration, *State Options Report* (Alexandria, VA: USDA FNA, updated March 5, 2026).

[57] Ariel Kennan and Sara Soka, *Benefit Eligibility Rules as Code: Reducing the Gap Between Policy and Service Delivery for the Safety Net* (Washington, DC: Beeck Center for Social Impact + Innovation, Digital Government Network, and Digital Benefits Network, 2022).

Step 4 is where federalism is most helpful. There are already working models, including the Department of Defense's shared conformance test service, which every service branch's systems have to pass (and keep passing) to be authorized to operate; the SNAP quality control rules mentioned above; and the IRS framework to which all commercial tax-filing tools must conform. We can do the same for all other benefit test suites that every state system has to pass. The state can (and should be required to) add its own test cases for its own options, and to confirm that state variations don't collide with federal requirements. For example, Minnesota's suite would add a scenario for its choice to use broad-based categorical eligibility, using the answer from its own state policies, while the federal suite still checks that the household's income is counted the way federal SNAP rules require.

Every state eligibility system is already built with federal match dollars and already requires an approved advance planning document (APD). Tie conformance to it: Federal match requires a passing federal suite and a published state suite, which also turns the APD from a review of a state's promises into a review of its results.

Why states may be more willing to do this now

Until recently, a state had limited financial reason to care whether its eligibility system implemented federal rules precisely. The error showed up in a sample, the sample produced a rate, and the rate produced some corrective action plans. That's over. Under OBBBA, starting in FY 2028, states with SNAP payment error rates at or above 6% pay up to 15% of their own benefit costs. For a large state, that's hundreds of millions of dollars a year, based on how accurately its systems apply rules that the state often can't clearly see or verify.

Documented criteria, reconciled definitions, authoritative sources, and conformance testing are the most direct routes to a lower error rate. For the first time, a state budget office has a direct, quantifiable financial stake in the things this paper recommends. That changes the conversation with states from "please participate in this federal initiative" to "here is how you avoid a nine-figure penalty," which we hope is a much shorter conversation.

Ultimately, federal dollars can be conditioned on adherence.

THE NEW ONE-STOP SHOP INTERFACES

"Anyone can now build an interface" is an example of a claim that sounds good in a policy paper, but has historically fallen apart. This was a promise of open data and open source software in government, as well, neither of which really materialized. But that changes if we do the work we've outlined in this paper, because it makes the hard part — the criteria, the tests, the authoritative sources, and the submission APIs — available for free for everyone to use.

Government still needs to provide baseline interfaces

AI doesn't mean the government should get out of the interface business entirely, at least not any time soon. **The government has a duty to serve everyone, including people without a smartphone or an AI assistant.** And now that anyone can quickly build a front door with AI, scammers will too: official-looking tools that steal data, charge unnecessary fees, and/or steal benefits from their users. Some people will still want one ".gov" door they know they can trust. **So it makes sense for government to provide a baseline set of interfaces: a website, an authenticated AI chat, a phone number to real humans, and in-person options.**

But that website and AI chat should be orders of magnitudes more affordable *and* higher-quality with a combination of the government-provided scaffolding outlined in this paper, and the advent of AI coding tools. We shouldn't need or have multiple competing government interfaces, and the ones we do have should be fast and affordable to build, still with great attention to usability and user needs. A small team with the tools outlined here could build an excellent USA.gov in a matter of weeks. Yet without the rules and tests, it can never be excellent, at any price tag.

Government shouldn't keep pouring its money and best people into the interface layer. **Government is the only entity that can say what the eligibility criteria are, reconcile them, authoritatively affirm facts about a person, resolve the cases the rules don't cover, and deliver the benefit. That's where the investment belongs.**

Government needs to support agentic benefits delivery

With published criteria and tests, any third party can build an interface that accurately assesses eligibility across federal and state programs, and submits applications on a person's behalf. Although this could include interested parties like nonprofit navigators looking to construct interfaces for specific populations or benefits, the most likely candidate today is an AI agent.

This is not a choice we get to make. Applicants are already using these tools at scale.[58] It won't be long before many people run their whole lives, bank accounts, restaurant reservations, and benefits enrollment, through whatever AI they've already handed their keys to. The websites we carefully user-tested and designed are not where applicants are going to turn.

[58] Chris Schmitz, Lewis Hammond, and Alan Chan, *Characterizing Agentic Flooding of Government Services* (Ithaca, NY: arXiv, arXiv:2608.16603, August 17, 2026).

Contact centers are already receiving calls from people who were given a confident wrong answer by a model that inferred eligibility rules from forum posts it found on the internet (not to mention receiving calls *from AI agents* seeking benefits on behalf of their human overlords). The question is not whether AI will mediate the relationship between people and benefits. It is whether the criteria those tools read are authoritative and testable.

An AI agent doesn't want the U.S. Web Design System or plain language. It wants rules. AI should do what it's actually good at: chatting with a person in their own language, figuring out which questions matter for them, and translating "I drive for DoorDash sometimes" into the facts the criteria ask for. The eligibility check it runs is ordinary, deterministic code built from government criteria, and has to pass the published test suite, into the government-run service APIs and government-run eligibility rules, to where the government ultimately (even if automatically) makes the decision. Government-provided scaffolding paired with AI can now generate something that was not possible before.

That division of labor also makes the AI verifiable. The AI is subject to the front-end tests, while the agency still makes the actual determination against its own authoritative data. Error responses (like a missing or invalid answer) can be sent back to and resolved by an AI agent, potentially much more effectively than a person can resolve the same error through a website.

ENFORCEMENT

Leverage executive branch authority to document eligibility criteria and reconcile them when possible

We've spent a lot of this paper explaining that agencies won't give up power voluntarily, so we need firm rules that enforce participation. The executive branch may be able to do more of this on its own than many think, given new precedents. This work should start boldly on the first day in office. Ultimately, Congress is likely needed to make the most substantial impact, and early collaboration and communication is a good start.

Whatever else the last several years demonstrated, the machinery of the executive branch turns out to be far more responsive to sustained, unambiguous direction from the top. A presidential directive to carry out the mission of this paper could spur significant action and progress. A new president should make this a condition of the job. Before offering anyone control of an agency, tell the candidate what the mission is: Document their agency's eligibility criteria, reconcile them with their peers, and cooperate with the agencies that hold facts their beneficiaries need. If a candidate won't commit to that, don't nominate them.

We can't think of anything an agency head should be *more* responsible for than delivering on its core services. That includes knowing the agency's own rules, reporting them accurately, keeping them current, and cooperating with the other agencies on which beneficiaries also depend. If an agency head refuses to carry out this work, that is a management failure, and it should be treated like one. Fire agency heads who don't follow this directive and screen for those who will.

Work with Congress to drive and compel reconciliation wherever necessary

Set up statutory requirements, alongside sticks and carrots

Some of these sticks and carrots exist in current law, while others may need Congress to expand or enforce them. An administration should immediately start using these levers to see how far the executive branch can get on its own, but communication and collaboration with Congress is encouraged.

Withhold funds. If an agency doesn't comply with laws passed by Congress, OMB can sit on its budget. For states, the lever is the federal money that pays for their eligibility systems and administrative costs. If a state unemployment agency doesn't increase the frequency of wage data collection and share wage data, it loses federal funding.

Withholding funds can work for vendors, too: If a state or federal benefits back-end system doesn't pass all the conformance tests, some or all of their contract dollars should be withheld until it does.

Reward cooperation. Sticks get compliance; money helps, too. Allocate funds that an agency or state can draw down to start this work early, like retiring an income verification contract before an upcoming renewal, standing up a conformance test suite, or migrating to an authoritative source. First movers get the money and the credit.

Shut the other doors. Once the income verification service is live, agencies may not collect or verify income separately. Alternatives are removed entirely and agencies that disobey face consequences.

The door has to be a *good* door, though. A mandatory service that doesn't work will produce shadow IT and workarounds. The agency that runs an authoritative source is on the hook for a published service level agreement (SLA) for uptime, response time, and accuracy; a named owner that the agencies depending on it can escalate to; and a change process where the people who query the service get a say and agreed-upon advance notice before it changes. If the source misses its SLA, the agency risks having the service (and the commensurate resources) reassigned. This will be a constant challenge, but if we can't do it well once, we definitely can't do it well in 100 separate places.

Establish a conformance safe harbor

Under this proposal, an agency that does the deconfliction work itself and publishes the result is essentially handing over the exhibit list for a future class action against its program. That's going to make it hard to get them to cooperate.

This is why the requirement has to be a mandate rather than an invitation. If publication is voluntary, no rational agency will go first, and any that does will be punished for it. That is the worst possible outcome, because it teaches everyone else to stay quiet.

But compelling an agency to produce evidence of its own noncompliance, and then leaving it fully liable for what it finds, is not fair. The mandate needs a companion: safe harbor legislation. An agency that publishes its criteria, maintains a test suite, discloses the inconsistencies it finds, and remediates any issues can raise that documented good-faith process as an affirmative defense.

The structure is not novel. Aviation has run on it since 1976: A pilot who reports a near-miss to NASA's Aviation Safety Reporting System within 10 days is protected from FAA enforcement, which is why the FAA has 50 years of near-misses to learn from. Medicine copied the model with the Patient Safety and Quality Improvement Act of 2005, which made reported adverse events privileged so hospitals *would* report them. Lucian Leape's argument for that system was that you can't fix errors you won't mention.

Safe harbor only works if the people it's meant for believe the line will be drawn fairly. Here, that means the line is written into statute, ahead of the publication requirement, rather than after it, not left to a future inspector general. The point is to make discovery safe, not to make errors free.

Similarly, once an established, government-run authoritative source exists, it should absorb the liability for any errors or issues. If not, and agencies and states are held liable for trusting its incorrect information, they will have every incentive to “double check” its work, creating the problem all over again.

Finally, the person on the receiving end needs safe harbor, too. Today when a system pays someone the wrong amount, that person is on the hook for paying it back. If a system has passed the test suite and the person gave true answers to the questions they were asked, the error should be the government's. SNAP already has a version of this in federal law (agency-error overpayments are collected at a lower rate and states can waive ones below a threshold). The agency should fix the rule and add the test case.

HOW TO DEFINE SUCCESS

Everything above assumes the White House, the agencies, the states, and Congress agree on what improvement looks like.

Many shared definitions will result in stricter requirements for some programs than their current design, while leading to more expansive access for others. Measured individually, program by program, reconciliation can look like a series of small losses and gains, and it's guaranteed that many program offices will complain they lost something. But measured across the whole ecosystem, it should clearly result in a net gain: fewer duplicative applications, faster answers, fewer wrong denials, less fraud, and a government that is finally capable of keeping its promises.

At the end of the day, there are three primary areas to make progress: the achievement of actual stated policy objectives (reducing homelessness, improving food security, making children healthier); improving efficiency and operations (lowering the cost to administer programs, reducing fraud, improving integrity); and delivering better experiences (improving uptake of eligible individuals, reducing burden and learning, compliance, and psychological costs).

To demonstrate progress, the team charged with this effort should identify a few core measures that can be both baselined (since we have poor data on the current state of affairs) and tracked along the way in each of these three areas. These measures should also anchor all parties to a shared North Star when the work gets hard.

Recent projects demonstrate what these measurements could look like. Amira's team is currently working to define measures that tell a holistic view of the "cost" to state budgets, across several domains (budget, beneficiary, operational, and outcome).[59] This stemmed from a research sprint to capture the "cost of the safety net" by tallying the program administration costs across SNAP, UI, TANF, and Medicaid, which posits that there may be states with better "cost to benefit" ratios.[60] These efforts preview the complexity of documenting *anything* consistently across the multiprogram, multistate, multilayer machinery that is benefits administration in the United States.

Rebecca Piazza, drawing on work convened under the 2021 executive order on customer experience, catalogs nine categories of measures people mean when they say "benefits should be better:" burden, timeliness, enrollment, benefit adequacy, administrative efficiency, program integrity, participant outcomes, equity, and customer experience. Each of those categories contains a dozen specific measures. Piazza identifies different ways those measures can generate tension:

- Directly: Improving one measure could lead to setbacks of another (e.g., more concierge, higher-capacity contact center agents that improve experience and access could increase the salary cost of contact center agents, and increasing uptake of benefits can increase total cost to program).

[59] Mat Poirot, *Resource for Practitioners: Cost of H.R.1 Implementation Framework Worksheet* (Washington, DC: New Practice Lab, September 30, 2026).
[60] Mat Poirot, *Resource for Practitioners: Cost of H.R.1 Implementation Framework Worksheet*.

- Indirectly: With limited staff capacity and topline budgets, making a choice to improve one piece could have an opportunity cost of not improving elsewhere.

There are most assuredly unintended consequences of selecting particular measures to focus on. SSA once counted customer "wait time" at field offices from the time you checked in at the kiosk, which did not capture the extensive time people waited *in the line* to check into the kiosk. Some government contact centers have optimized their "wait time" by excluding calls that opted for a call back, but may never have successfully connected with an agent. In the "cost of safety net" project, we learned that integrating forms did not inherently result in faster processing times – again, because the hard work that would result in that outcome measure change was not undertaken.

This is where a North Star "outcome" focus can be a great starting place:

- No person should have to retell the government information, if it already has it and the individual wants them to use it.
- Government shouldn't do work twice (or more than twice); if it is adjudicating you for a benefit, and you want it to, it should adjudicate you for *everything* for which you are eligible.
- We should be getting more value for every taxpayer dollar we spend.
- When someone needs help fast, they should get it (depending on context, we could define this as "within a week," like the "Facing a Financial Shock" North Star).

Somebody has to own those outcomes: one person or office with the authority to make agencies move, measured on the whole-system result and not on any one program's. Underneath that, a handful of measures across budget, operations, beneficiary experience, and outcomes shows whether the work is actually paying off.

In a future benefits integration effort, measures could include:

Budget

- Lower cost per case to process
- Reduced vendor spend on the same data element

Operations

- Reduced "requests for information" seeking additional information, cases that sit in purgatory
- Reduction in "churn" (repeat work of agencies processing new applications from previously enrolled individuals who faced barriers to recertification)
- Conformance (the share of implementing systems currently passing the published test suite. Since this testing framework doesn't exist today, we can't measure it)

Beneficiary Experience

- Percentage of eligible people enrolled (this shows whether the process is accessible)
- Time to apply (if we're not in an "opt-out" system and people need to apply)
- Time to benefit receipt (could be first payment, when your health insurance is active)
- Percentage of procedural denials (this separates "you do not qualify" from "we could not manage to find out." The second is an administrative failure, and we want fewer of those.)
- Ex parte and auto-renewal rates (this indicates whether the shared verification layer is actually being used, and whether we are getting more proactive rather than merely faster)

Outcomes

- Percentage of veterans who are homeless
- Percentage of Americans covered by health insurance
- Percentage of children in America experiencing hunger

If a president wanted one line for the State of the Union, it's this: No one should have to tell the government the same thing twice (and ideally, not even once!) to get what they need.

WHAT WE CAN DO NOW

Fully climbing the rungs of this ladder will require action from both Congress and an executive branch that are motivated to get things done and use their resources to do so. As we've outlined, these are not actions that will one day evolve purely from experience and increased maturity; they require political will.

To prepare for this future, a number of efforts are underway and can be built upon. These deliverables can be produced in the absence of power, by entities outside of government:

- Building the template for documenting existing benefit program eligibility requirements, and developing a schema to handle the inherent complexity of different eligibility rules across states and criteria that may be dynamic or use multiple sources for verification
- With this documentation, mapping existing definitions across programs, and beginning to build out a plan for reconciliation that can occur based on where criteria or verification standards are codified and the impact of making each change
- Building working authoritative sources in pilot states, starting with income verification
- Encouraging agencies to flag definitional conflicts in their technical assistance on every benefits bill under consideration, so Congress hears about them before the vote
- Preparing congressional offices for the work that will be required of them, including briefing members on the amount of funds their state is losing now on duplicative efforts, and how their state currently compares on administrative cost spending (spoiler: everyone has room for improvement and savings)
- Shaping a clear vision of what we want: how we proactively get people what they need, but maintain important privacy controls and individual ownership of data

Government should reconcile program rules where nothing justifies the difference, be the authoritative source for the criterion data it already holds, and prove that its systems can apply both sets of information correctly. If this is done well, Americans will no longer need to jump between government sites and applications to figure out what support they're eligible for. One submission — or possibly none at all — will be enough.

Appendix A: The Bold Future of Integrated Benefits

A bolder, more ambitious horizon is also possible. One cautionary note before we share this list: Everything below is a structural bet about where benefits should live and who should run them. We may decide differently as this work continues. But whether programs get consolidated, centralized, or stay exactly the same, government still needs a way to document rules, test implementations, measure outcomes … *and* to deliberately change operating models when called for. That last capability is worth its own study, and it's reusable far beyond benefits.

Some bold ideas could look like:

A Department of Government Services. Does it make any sense to have this many benefit lines in this many places? No.

The chart of the American safety net is not a design, rather an archaeological record. SNAP sits at USDA because it started in 1939 as a way to dispose of agricultural surplus and prop up farm prices, and school meals are there for the same reason. Obviously, the Low-Income Home Energy Assistance Program lies in HHS. Programs often live wherever the coalition that passed it happened to have a friendly committee and a willing secretary. Nobody would build it this way starting over, and nobody *did* build it this way — it accreted, program by program, rule by rule.

Could the delivery and updating of safety net benefits, and even some administrative services (like renewing a passport or getting a camping permit in a national park), be significantly consolidated and run by one very small team that could sit around two pizzas? Why not?

Now, absorbing programs is a wholly different animal from reconciling definitions or operating service intake requests: It means moving money, staff, systems, and constituencies. SSA is the closest working model we have to multiple program service delivery directly with the American public. The fact that it runs retirement, disability, SSI, and Medicare enrollment federally, at enormous scale, suggests that it is possible. This also took four separate acts of Congress across 37 years, each *adding* a program rather than consolidating one, which says something about how fast this happens. People are already talking about a Department of Services that at a minimum could establish and enforce service delivery standards.[61]

Tax as foundational infrastructure. The IRS could become the foundational layer for verifying income, adjudicating eligibility, and delivering benefits. This would require a shift in perceiving the IRS as shared infrastructure rather than (solely) as a tax collection agency. For example, the IRS could determine eligibility for Medicaid.

The IRS already has the tools to do this. It proved it, on short notice, with the Advance Child Tax Credit in 2021 and the pandemic Economic Impact Payments. It delivered money into most Americans' accounts, with no application, no office, no interview, and no thousand-screen

[61] Amira Boland and Tara McGuinness, *Reinvention Blueprint No. 1: A Government That Works for People* (Just Security, August 11, 2026).

eligibility system in between. As mentioned earlier, the Canadian tax form *is* a benefits form. Consent to file your taxes in Canada is consent to get some cash back (if you are eligible).

One could decide the tax account is how *all* federal benefits get distributed. It combines the citizenry social contract: You pay taxes, and you get things, too. While the IRS is not where the expertise for reimbursing for Medicaid services or managing EBT vendors lives, it could at least be the interface for enrollment.

Instant adjudication. Once every criterion has one authoritative source and every rule has a test suite, there's no good reason a determination should take longer than a credit card swipe. Today SNAP takes 30 days to decide, Medicaid takes 45, unemployment insurance considers itself on time if your first check shows up within three weeks, and VA disability claims routinely take months. Marina knows it's possible for many VA disability claim decisions to be instant, because she built a working prototype to do it in 2015. (It's a longer story as to why it was politically untenable to launch.)

This doesn't have to be all or nothing for a given benefit, but when the process can be tweaked such that some subset of applicants can have their eligibility verified and benefits delivered instantly, we should do that.

Population-level or single-fact eligibility as the default. Before designing a new, complicated formula for determining eligibility, the first question should be: Can someone be deemed eligible based on one fact (like participation in another program, presence of a health condition, geography, or another observable characteristic)? How might we design this program to require no new or individual data collection? For example, during the pandemic, USDA issued nationwide waivers that made entire schools (based on their location) eligible for free lunch. Attaching benefits universally to children has been used with cash benefits like the Child Tax Credit and Trump Account deposits.

Centralize more programs. States — or worse, counties — individually administering federal programs like TANF or UI generates massive inefficiency. Is it fair that residents of one state have a worse experience or less access than someone in another state? We have precedent for undoing poorly administered decentralized efforts: In 1972, Congress federalized three state-administered adult assistance programs (Old Age Assistance, Aid to the Blind, and Aid to the Permanently and Totally Disabled) into SSI.

A CBO scoring process that penalizes bespoke benefits. Going forward, new legislation should score as more expensive when it generates new definitions that deviate from existing ones, adds new requirements to existing processes, requires different verification standards than existing determinations, doesn't demonstrate how the minimum-viable, single-fact path was pursued first, requires new technical infrastructure to implement, or doesn't even default to an opt-out (versus opt-in) model.

Appendix B: What Success Looks Like

Recall Dana from our earlier scenario; now let's run her household through the world we're imagining. Dana was laid off at the end of February; her husband, Marcus, earns $1,900 a month; their daughter, Amaya, 16, makes about $380 a month at a part-time job; their son, Theo, is 8; Dana's mother, Ruth, 68, lives with them on $1,240 a month in Social Security and cooks for herself; Marcus' brother sends $250 a month toward the rent; and Dana sold some stock in January.

One night in the car, Dana opens her favorite AI assistant to plan dinner, then thinks to ask it for help applying for potential benefits she's eligible for now that she's lost her job.

She certifies who she is with a government-certified credential, which she has previously linked with her AI assistant. She only has to do that once — no mailing a photocopy of her driver's license or showing up in person.

She's never asked for her income. Her employer reported every paycheck as it was paid, and so did her husband's, so the government already knows their income. Nobody asks Dana to upload pay stubs or find last year's tax return.

She answers a few questions about her household — who lives with her, who buys and prepares food together — once. These are things only Dana knows, and policy accepts her word as fact.

Facts the government already holds are never requested or even mentioned. Citizenship is confirmed once against the authoritative source. Amaya's age, Theo's school enrollment, and Ruth's Medicare status are not questions Dana has to answer or documents she has to produce; the government already knows these facts.

But there's a snag. Dana was paid early in February because payday fell on a holiday, so two paychecks land in one SNAP counting period and she looks, at first glance, like she earned twice as much. Instead of receiving an instant answer about SNAP, she has to wait for a human to take a look. The system flags her claim internally *before* denying her, because that's the protocol. A caseworker with actual authority immediately sees the issue and reallocates the payment to the right period in SNAP's eligibility system. The wage record at the source is untouched (because it was not *wrong*; the issue is on the SNAP side). That scenario soon becomes a new SNAP system automation, so the next person paid early on a holiday is automatically approved.

Dana gets answers with explanations. She qualifies for SNAP at a specific amount. She does not qualify for the heating program, because her state set its threshold at 150% of the poverty guideline and her household is at 162%. Her application for unemployment benefits is approved right away. (Under today's UI systems that would be a miracle; with per-paycheck reporting in place, it's just Tuesday,) She's also cautioned about a consideration she didn't ask about: Her AI notices that Marcus has overtime on offer in March, and suggests that if he takes it, the household income will cross a cliff that costs more in lost benefits than the overtime will pay.

Dana had no idea. (Almost nobody does.) That last response is possible only because someone did the hard work of mapping the incompatibilities across programs.

The already approved money and in-kind support arrives the next day, to her payment account already on file.

Dana used a one-stop shop, but nobody procured it, and no program office staffed it. There was no launch event, no press release, and no 2027 modernization initiative. It exists because of what lies underneath.

Dana still had to decide she wants help and go looking — not a failure of the design, but the nature of a benefit that depends on a fact only she can report (being newly unemployed — though in time, we could largely automate this news, too). She still had to answer the questions only she could answer. But this abolishes the "time tax" — the part of administration that consists of proving the same thing over and over to a system that already had access the whole time.[62]

[62] Annie Lowrey, *The Time Tax: How the Government Wastes Our Time—and How to Fix It* (New York, NY: Ecco, 2026).

Appendix C: Open Questions

To move toward this future, we also need to address these questions:

- How is consent managed? When does a citizen affirm they *want* to be considered for benefits for which they are eligible, or do they even need to? What would the demand look like for opting out?
- How are errors corrected? When a fact is wrong, someone with power has to fix it at the authoritative source, and everyone who relied on it previously has to get the update. At the VA, it was a not very funny story that malicious ex-spouses would call to report their former partners as dead, and changing them back to “alive” was very hard.
- How do you manage exceptions and emergencies? If someone is incorrectly locked out of all benefits at once, because they are now connected, how do we ensure recourse at speed?
- How do we improve the quality and availability of real-time data for all types of people? Per-paycheck wage reporting works great for W2 employment but doesn't fully cover the 1 in 10 Americans who are self-employed, or many others who manage income across different sources. If a system is more centralized federally, how do you still enable connection to more regional sources like credit unions or smaller employers?
- What is a manageable and ambitious scope of what the “safety net” is for the first passes of integration? In this paper, we’ve named a few lighthouse places where this at a minimum must begin — like SSA, USDA, HHS, HUD, DOL, IRS, and VA — but even within those is a complex web of services.
- How do we roll out integrated benefits in a way that designs for access and concerns of many different customers? The unbanked are a population that is rightly often raised as requiring exceptions in program delivery; however, rather than solving one-off fixes, how do we use this as an opportunity to better connect them to banking?
- How will Congress support consolidated or integrated use of appropriations for shared infrastructure? The current color of money extends down to the way in which identity, verification, and use of authoritative sources are in some cases individual line items inside each participating program budget. What if one agency needs to bear the cost of development in its budget, but that means two other agencies don’t need to spend a dime?

Appendix D: Citations for the 30 Years of One-Stop Shops Table

Citations on 30 Years of One-Stop Shops:

- National Performance Review and Government Information Technology Services Board, *Access America: Reengineering Through Information Technology* (Washington, DC: National Performance Review, February 3, 1997)
- The White House, *Technology Access and Innovation* (Washington, DC: The White House, June 2000).
- The White House, *President Launches FirstGov: A Single, Easy-To-Use Website for Government Services* (Washington, DC: The White House, September 22, 2000).
- U.S. House, *E-Government Act of 2002*, H.R. 2458, 107th Cong. (2001) (introduced by Rep. Turner) (enacted as Public Law 107-347).
- Office of Management and Budget, *Report to Congress on the Benefits of the E-Government Initiatives* (Washington, DC: Office of Management and Budget, January 2009).
- Cindi Fukami and Donald J. McCubbrey, *Colorado Benefits Management System (C): Seven Years of Failure*, Communications of the Association for Information Systems 29, art. 5 (2011).
- U.S. Department of Labor, *GovBenefits.gov Becomes Benefits.gov* (PR Newswire, September 30, 2010).
- The White House, Office of the Press Secretary, *White House Launches Presidential Innovation Fellows Program* (Washington, DC: The White House, August 23, 2012).
- Leslie Small, *Oregon, Oracle Settlement Ends Fight over Failed State Exchange* (Fierce Healthcare, September 19, 2016).
- Billy Mitchell, *VA Launches Rudimentary Vets.gov on Veterans Day* (FedScoop, November 11, 2015).
- The Associated Press, *Deloitte Apologizes for Botched Rhode Island Benefits System* (The Associated Press, April 13, 2018).
- Executive Order 14058, *Transforming Federal Customer Experience and Service Delivery To Rebuild Trust in Government*, 86 Fed. Reg. 71357 (December 16, 2021).
- U.S. General Services Administration, *Department of Labor, General Services Administration to Centralize Government Benefits Information on USA.gov and USAGov en Español* (Washington, DC: U.S. General Services Administration, August 7, 2024).
- Russell T. Vought, *Scaling Use of Login.gov to Deliver a Universal Sign-on for Public Services*, M-26-18 (Washington, DC: Office of Management and Budget, August 31, 2026).
- Executive Order 14432, *Streamlining Access to Government Services Through America.gov* (September 29, 2026).

Appendix E: Citations for the What's Actually Worked Table

Citations on What's Actually Worked:

- Social Security Administration, *When to Sign Up for Medicare* (Baltimore, MD: Social Security Administration, updated June 10, 2025).
- Social Security Administration, *List of State Medicaid Programs for the Aged, Blind and Disabled*, POMS SI 01715.020 (Baltimore, MD: Social Security Administration, effective August 2, 2016).
- U.S. Department of Agriculture, Food and Nutrition Administration, *Summer Food Service Program History* (Alexandria, VA: USDA FNA, updated March 31, 2019).
- U.S. Senate, Committee on Agriculture, Nutrition, and Forestry, *The Child Nutrition and WIC Reauthorization Act of 2004*, S. Rep. No. 108-265 (Washington, DC: U.S. Senate, June 7, 2004); U.S. Department of Agriculture, Food and Nutrition Service, *Direct Certification in the National School Lunch Program: State Implementation Progress Report to Congress, School Year 2017-2018 & School Year 2018-2019* (Alexandria, VA: USDA FNS, June 2021).
- Social Security Administration, *Medicare Part D Extra Help (Low-Income Subsidy or LIS)*, POMS HI 03001.005 (Baltimore, MD: Social Security Administration, effective April 9, 2024).
- U.S. General Services Administration, Office of Evaluation Sciences, *Increasing Servicemember On-Base TSP Enrollment* (Washington, DC: U.S. General Services Administration, 2015); Thrift Savings Plan, *Revision to Implementation of the Blended Retirement System*, TSP Bulletin 17-U-3 (Washington, DC: Thrift Savings Plan, August 29, 2017).
- U.S. Department of Agriculture, Food and Nutrition Service, *Child Nutrition Programs: Community Eligibility Provision—Increasing Options for Schools*, 88 Fed. Reg. 65778 (September 26, 2023).
- U.S. Department of Health and Human Services, Centers for Medicare & Medicaid Services, *March 2025: Medicaid and CHIP Eligibility Operations and Enrollment Snapshot* (Baltimore, MD: CMS, June 27, 2025).
- Internal Revenue Service, *An Overview of the 2021 Advance Child Tax Credit*, Publication 5534 (Washington, DC: Internal Revenue Service, October 2021).
- New Jersey Department of the Treasury, Division of Taxation, *Affordable New Jersey Communities for Homeowners and Renters (ANCHOR)* (Trenton, NJ: New Jersey Division of Taxation, updated September 16, 2026).
- U.S. Department of Agriculture, Food and Nutrition Administration, *Summer EBT Household* (Alexandria, VA: USDA FNA, updated July 22, 2026).
- U.S. Department of the Treasury, *Treasury Announces the Completion of Automatic Enrollment Today for Trump Accounts* (Washington, DC: U.S. Department of the Treasury, October 1, 2026); Jessica Dickler, *Trump Account Holders May Receive $250 Dell Grant Starting Monday, Dell Foundation Says* (CNBC, August 31, 2026).