Triangulating Across U.S. Federal AI Transparency Regimes

Emma Lurie[1], Emma Fauser[2], Qing He[1], Danaé Metaxa[1], and Sorelle A. Friedler[2]

[1]University of Pennsylvania, Department of Computer and Information Science, Philadelphia, PA, USA
[2]Haverford College, Department of Computer Science, Haverford, PA, USA
ewlurie@seas.upenn.edu, efauser@haverford.edu, qingh@design.upenn.edu, metaxa@seas.upenn.edu, sorelle@cs.haverford.edu

Abstract

Federal AI systems can deny benefits or flag individuals for deportation, but the public disclosures meant to make those systems visible are fragmented and unevenly detailed. This paper examines three existing U.S. federal transparency regimes—System of Records Notices (SORNs), Information Collection Requests (ICRs), and the AI Use Case Inventory—and asks how well they, individually and together, describe government AI use. We find that no single regime fully reveals how the government constructs or deploys AI: each discloses different aspects of a system, and the current disclosure infrastructure makes it very challenging for the public to track specific AI systems across regulatory regimes and over time. Persistent identifiers are absent, granularity varies widely, and the annual AI Use Case Inventory cycle means federal agencies can deploy systems months before appearing in any official record. Using hand-validated zero-shot classification and cross-document entity resolution, we contribute a triangulation method that links disclosures across all three regimes and present two case studies. Our case studies finds that linking records provides greater insight into government AI use, but even linked records would constitute insufficient oversight compared to what public reporting has revealed about the same systems. We trace each regime's disclosure weaknesses to its original administrative purpose, showing these gaps are structural, and offer recommendations focused on the AI Use Case Inventory as the mechanism best suited for public-facing transparency: (1) a broad and consistently applied AI system definition, (2) persistent system identifiers with cross-references to related disclosures, and (3) restored public visibility into risk management processes.

Website — https://govAIfiles.org

Code and data —
https://github.com/triphora/govaifiles.org

Extended version — https://arxiv.org/abs/2607.29540

1 Introduction

In December of 2025, a U.S. citizen was walking home when he was stopped and detained by U.S. federal immigration officials. Agents handcuffed him and scanned his face using a phone app; he was only released when the app confirmed his citizenship. Reporting revealed that agents likely used Mobile Fortify, a facial recognition and fingerprint-matching app used by Immigration and Customs Enforcement (ICE) to determine a subject's citizenship status.

The app, first reported by 404 Media (Cox and Cleek 2025), raised a number of questions and concerns about AI governance and accountability in a U.S. federal context: was this app being used as directed, were there avenues for redress and oversight, and was its use even legal? Proposals for answering these and other AI governance concerns often rely on transparency mechanisms (White House Office of Science and Technology Policy 2022; OECD 2019; Kingsman et al. 2022). But new transparency mechanisms may or may not be needed; the U.S. federal government already operates under disclosure requirements, including some designed specifically for AI and others that may disclose government AI use related to databases and data collection.

In this paper, we identify three existing U.S. federal government transparency regimes (further detailed in Section 3) and consider the extent to which they make AI applications visible to the public. These existing regimes are:

1. **System of Records Notices (SORNs)** are required notices whenever a federal agency creates or changes the functionality of a database containing personal information. Designed in the 1970s with a focus on government data storage, SORNs may now reveal information about AI systems that make use of that data.
2. **Information Collection Requests (ICRs)** comes out of the Paperwork Reduction Act and requires agencies to report on any data collection processes, whether data is collected by paper or electronically; when agencies collect data for an AI system, this may be revealed via ICRs and supplementary documents such as Privacy Impact Assessments (PIAs).
3. Finally, the federal **AI Use Case Inventory** is a recent effort specifically tailored to make AI use by the federal government transparent. All Executive Branch agencies are required by law to disclose all AI systems that are in use by their agency to the Office of Management and Budget (OMB) and disclosed to the public.

Our research considers what existing federal AI disclosure regimes reveal and obscure about government AI use, and what cross-regime comparisons illuminate that single-regime analyses cannot. Linking records across

regimes is a substantial methodological challenge: the three regimes share no common system identifiers, were not designed to be interoperable, and each uses vocabulary shaped by its own administrative logic. Using zero-shot learning and cross-document entity resolution techniques developed to bridge disparate text corpora, we link AI deployments across all three transparency regimes.

We find that the triangulation of AI regulatory regimes reveals insights that no individual transparency regime provides. Using two high-impact systems as case studies—*Automated Targeting System* and *Mobile Fortify*—we demonstrate that connecting multiple transparency regimes allows stakeholders to better understand the assemblage of systems, data sources, models, and oversight that comprise the entirety of the system. We find that it is currently difficult for the public to track a specific AI system across regulatory disclosures and over time: persistent identifiers are absent, between 80 and 97 percent of records in each corpus have no confirmed cross-regime link, and the annual AI Use Case inventory cycle means systems can be deployed for months before appearing in any official record. Further linkage and transparency concerns include vocabulary divergence, inconsistency in the amount of desired and revealed detail, and mission creep, such that a system use changes without requiring the associated disclosure to be updated.

We close with policy recommendations for improving government AI transparency. The AI Use Case Inventory is a promising mechanism, but several fixes would improve its efficacy towards public transparency: removing the "principal basis" language that limits the AI definition, adopting persistent system identifiers with cross-references to other government filings, and restoring public visibility into risk management processes. We can reform the existing disclosure ecosystem without whole cloth reinvention.

Contributions

In summary, this paper contributes:

1. **Evidence that triangulation adds value**: using *Automated Targeting System* and *Mobile Fortify* as case studies, we show that cross-regime analysis reveals more about high-impact AI systems than any single transparency regime currently can.
2. **A characterization of systemic gaps in federal AI transparency**: we identify three recurring failure modes — timing mismatches, inconsistent granularity, and mission creep — and trace each to the history and administrative logic of the underlying mechanism.
3. **A public transparency tool for navigating the AI Use Case Inventory**: We develop https://govAIfiles.org a user-friendly web application that makes it easier for the public to search and explore government AI use.
4. **A cross-regime linking methodology**: we develop a zero-shot learning and cross-document entity resolution method to connect AI disclosures across three federal transparency regimes not designed to interoperate.
5. **Policy recommendations for improving federal AI transparency**: we offer targeted recommendations for the AI Use Case Inventory, including clarification to the definition of covered AI systems, the adoption of persistent system identifiers, and reinstating robust risk management disclosures.

2 Literature Review

The role of transparency in AI governance

Transparency is widely recognized as foundational to AI governance (Jobin, Ienca, and Vayena 2019). When government entities deploy AI systems, individuals can rarely opt out of being subject to those systems, making transparency an important mechanism for correcting the informational and power asymmetries between governments and constituents (Hickok 2024). Transparency is a prerequisite for rights-preserving AI governance: without the ability to know how, why, and to what end AI systems are deployed, other accountability mechanisms are undermined (AlgorithmWatch and Bertelsmann Stiftung 2020).

Yet disclosure alone does not guarantee meaningful transparency. Ananny and Crawford's (Ananny and Crawford 2018) concept of "seeing without knowing" explains that having knowledge of a system, is not the same as knowing how it works or being able to govern it. Often, efforts at various forms of transparency or explainability have actually been found to be directed at assessing specific properties of a system, such as its potential for discrimination, contestation, or recourse (Selbst and Barocas 2018; Ustun, Spangher, and Liu 2019). Algorithmic impact assessments are a common operationalization that assesses the achievement of normative goals, such as non-discrimination, via transparent reports that document the process and results of system tests (Reisman et al. 2018) and that have been proposed as part of governance processes in the vein of environmental impact assessments (Selbst 2017). Algorithmic impact assessments form an increasing part of governance approaches to AI (Selbst 2021), and there are many methodological concerns about how best to use these instruments as effective governing tools (Watkins et al. 2021) especially with regards to including marginalized communities in the determination of what impacts are measured (Metcalf et al. 2021). The AI use case inventory we study brings some of these concerns to the fore—impact assessments are required as part of the risk management disclosures for high-impact AI systems.

Studies of Government AI Use

The public record on how artificial intelligence systems are built, funded, and overseen by the U.S. federal government is scattered across government websites and difficult to piece together into a coherent narrative. Gaps or fragmentation in disclosure translate directly into gaps in oversight, leaving affected communities with limited recourse. Scholars have approached this problem by examining different government datasets. Wang et al. (2022) show that ICE constructed a domestic surveillance infrastructure through procurement relationships that largely bypassed public scrutiny, drawing on data from FOIA requests and procurement records that had previously generated little public attention. Levendowski (2021) locates transparency in federal trademark filings. Because the USPTO requires evidence of use, trademark reg-

istrations can compel private surveillance companies to disclose operational details about products that other regulatory channels (FOIA, patents) tend to obscure. Bateyko and Levy (2025) identify a different overlooked site of AI governance: federal grant notices. Agencies directing billions in discretionary funding shape how state and local entities develop and deploy AI, yet this grantmaking activity has developed largely outside the transparency norms that govern procurement. Across these three bodies of work, a pattern holds: consequential AI governance activity occurs in administrative registers that were not designed with AI transparency as a goal, and are rarely analyzed together.

What remains underexplored is whether existing federal disclosure regimes, examined in combination, can yield meaningful public insight into high-impact AI systems, and what cross-regime comparison reveals that any single regime cannot. We examine three such mechanisms (System of Records Notices, Information Collection Requests, and the AI Use Case Inventory), tracing how each renders AI systems visible or invisible, and where their coverage diverges. The aim is not to call for new regulatory infrastructure wholesale, but to identify targeted reforms grounded in the administrative history and logics of what already exists.

3 Existing Transparency Regimes and AI

Debates about government surveillance, administrative burden, and public accountability for automated systems and data collection are not unique to the AI era (Ruggles and Magnuson 2023). They have produced a layered set of disclosure regimes stretching back to at least the 1970s (U.S. Department of Health, Education, and Welfare 1973). In this section, we identify these existing regimes and describe how they could help us gain visibility into AI system use.

System of Records Notices

System of Records Notices emerged from 1970s concerns about government computerization and its implications for individual privacy. In drafting the Privacy Act of 1974, Congress relied on an advisory committee report documenting the risks presented by the increasing use of electronic information technologies, which had begun replacing traditional paper-based government systems (U.S. Department of Health, Education, and Welfare 1973). To address these risks, the report recommended five Fair Information Practice Principles: openness, individual access, collection limitation, use and disclosure limitations, and information management (U.S. Department of Health, Education, and Welfare 1973). These principles remain at the core of privacy legislation at the federal, state, and international level (Gellman 2025; Hartzog 2017).

The Privacy Act operationalized these principles by requiring federal agencies to publish a System of Records Notice in the Federal Register whenever an agency creates or modifies a system containing personal information about individuals. These notices must include the system's name and location, the types of individuals and records included, how the information will be routinely used and by whom, the agency's policies on data storage and retention, and procedures for individuals to access and challenge their records (United States Congress 1974). Thus, AI systems can be revealed when included as part of the description of how the data will be used, and when identified as the data underlying an AI system, the description of the data can shed light on how such systems work.

Information Collection Requests

The Paperwork Reduction Act (PRA) regulates information-collection burdens imposed by the federal government. Enacted in 1980 and revised in 1986 and 1995, it requires agencies to obtain OIRA approval before collecting information from ten or more members of the public (Shapiro 2012, 2020). The review process requires agencies to justify the need for the collection, estimate the burden imposed on citizens, provide opportunities for public comment, and display a valid OMB control number (Shapiro 2012; Levy 1994). Notice of the collection must be filed with the Federal Register. In cases where personally identifiable information (PII) is collected, agencies are also required to file a SORN before filing their ICR (United States Office of Personnel Management 2011). OIRA's goal is to minimize federal information-collection burden while also maximizing the utility of the information collected. More broadly, OIRA acts an institutional "information aggregator" that collects dispersed information from agencies, the Executive Office, state and local governments, and the public, and as a "guardian of a well-functioning administrative process" (Sunstein 2013). It is this aggregation feature that we take advantage of in this study as we analyze the documentation included in Information Collection Requests (ICRs) and posted to a federal government website.

The goal of the PRA was to respond to concerns over the cost burdens, often in terms of time, imposed on citizens, state and local governments, and businesses by federal paperwork (Levy 1994; Samaha 2015). Federal agencies collect a large amount of information from the public and retain substantial discretion over the form and burden of that collection. The PRA's aim was to force agencies to screen out duplicative, poorly designed, or unnecessary collections (Levy 1994). The statute's performance is often criticized. On aggregate burden reduction, there is limited evidence that the PRA has achieved either large-scale burden reduction or accurate burden estimation (Shapiro 2012, 2020; Pahlka 2023; Samaha 2015). Yet, OIRA review improves some individual collections and may deter some ill-considered collections from ever being proposed (Shapiro 2012, 2020). The PRA is therefore best understood as a procedural quality-control regime with uneven aggregate effects: it improves some collections and creates accountability but has not reliably reduced the cumulative paperwork burden it was designed to control.

Importantly for our purposes, agencies file an ICR with OIRA before manually or automatically collecting digital or biometric information as well as more traditional paperwork. The result is that AI systems built on such collected data, such as facial recognition systems, can be found described in these filings.

AI Use Case Inventory

Unlike the regimes that preceded it, the AI Use Case Inventory was designed explicitly around AI. The inventory was originally required under the first Trump Administration via an executive order on trustworthy AI (EO 13960), and later codified in law under the Advancing American AI Act. The Biden Administration continued this implementation, and provided associated guidance on risk management through an OMB management memorandum (M-24-10 (Office of Management and Budget 2024)) which was updated by the second Trump Administration (M-25-21 (Office of Management and Budget 2025)), though remained substantively the same in many ways (Friedler and Selbst 2025).

Based on reporting guidance from OMB associated with these memoranda, agencies must annually report publicly and to OMB's Office of the Federal CIO (OFCIO) each AI system used by the agency, whether developed in-house or acquired for use, with exceptions for research systems, national security systems, and a few other uses. Each reported use case must include basic information about that use case, such as its purpose and the type of AI system, as well as risk management information in the case of high-impact systems. High-impact systems are identified via a specific definition in the OMB management memoranda and have an associated list of specific systems that are presumed to be high-impact (Office of Management and Budget 2024, 2025). These include risk assessment systems related to immigration, biometrics used in public spaces, vehicle tracking of non-governmental vehicles in a law enforcement context, patient diagnosis systems, and other immigration, law enforcement, healthcare, government services, and safety or rights related AI system uses.

4 Methods for Triangulating Across Regimes

In order to develop a dataset on which we can conduct a cross-regime analysis, we first collect and filter the data to include only identified AI-related entries and standardize the resulting information. Then we create methods for triangulating across these collected AI-related entries.

Identifying and Standardizing AI-related Entries

SORNs We identified AI-relevant Systems of Records Notices (SORNs) using a zero-shot large language model classification workflow applied to a corpus of federal SORN JSON files. The underlying corpus consisted of 3,735 SORNs spanning 2010-2025 pulled from the Federal Register API. We then classified these SORNs as AI-related using zero-shot classification on GPT-5.4 mini. Model outputs were reviewed on a random sample of 122 SORNs flagged as AI-related by the classifier, of which 85 (69.7%) were confirmed and 37 (30.3%) were incorrectly labeled as AI-related. The final prompt (see Extended version Appendix), distinguished systems that perform matching, scoring, recognition, prediction, ranking, approval or denial support, risk assessment, or other automated analytical judgments from systems that primarily store, retrieve, exchange, or display records. It also returned a binary `is-ai-related` flag indicating whether machine learning or AI was explicitly stated in the notice, short quoted evidence phrases copied from the SORN text, and brief reasoning. Model outputs were then reviewed as part of the evaluation process.

ICRs We collected all Information Collection Requests (ICRs) received by the Office of Information and Regulatory Affairs (OIRA) from January 1, 2018 through March 31, 2026, yielding 37,871 ICR records. We scraped ICR listings and detail pages from reginfo.gov; each ICR's detail page was parsed to extract structured metadata including the submitting agency, ICR reference number, title, abstract, and links to supporting documents.

To identify AI-relevant materials within this corpus, we applied a filter to identify OMB supporting statements, Privacy Impact Assessment (PIA), Privacy Threshold Analysis (PTA), System of Records Notice (SORN), Algorithmic Impact Assessment (AIA), or civil liberties assessments; as well as documents that contained a predefined set of AI-related vocabulary (see Extended version Appendix for the filter procedure and AI-related vocabulary). After filtering, 3,378 ICRs moved onto the classification stage.

Candidate documents were classified using the GPT-5.4-mini model. Each document was submitted with a structured prompt (see Extended version Appendix) asking the model to assess whether the described information collection involved AI. Classification was applied at the document level; ICRs with at least one AI-related document were flagged as AI-related at the ICR level. We manually reviewed 100 of the ICRs that were flagged as AI matches, and found 75% to be AI-related. The full batch produced 1,146 AI-related ICRs (33.9% of the 3,378 candidates). See Extended version Appendix for further details.

AI Use Case Inventory We analyzed all federal agency AI Use Case Inventory disclosures (2023-2025). To support cross-agency and cross-year comparison, we normalized column names, response categories, and standardized agency and bureau names to canonical forms using a crosswalk and fuzzy-matching procedure (see Extended version Appendix for details). Because all entries in the AI Use Case Inventory are AI technologies, no classification was required.

Cross Regime Analysis

A methodological challenge in this paper is determining when records from the three regulatory regimes (SORNs, ICRs, and AI Use Case Inventory) refer to the same underlying AI system. These documents comply with different regulatory requirements, focus on different elements of AI systems, use different vocabularies for important terms, are filed by different offices at different times, and rarely cross-reference one another.[1] The task of linking them is an instance of *cross-document entity resolution* (Beheshti et al. 2017): identifying when textually dissimilar records

[1]While ICRs that involve PII require an agency to file a SORN first, they sometimes still may not include the associated SORN or SORN identifier directly in the ICR.

	AI Use Case Inventory	SORN	ICR
Coverage	FY2023, FY2024, FY2025	2010–2025 (Federal Register)	Jan 2018–Mar 2026 (OIRA)
Collection unit	Use case record, per agency	Published notice	Information collection request
Total collected	710 (2023); 2,071 (2024); 3,542 (2025)	3,735 notices with full text	37,871 ICRs; 31,670 with downloaded docs
AI identification method	Agency reported	Zero-shot LLM classification	Two-stage keyword filter + zero-shot LLM
AI-related records	All records (reported by agency)	554 of 3,735 (15%)	1,146 of 3,378 classified ICRs (34%)[a]

Table 1: Dataset collection and AI-related findings by disclosure regime.
[a] The 34% rate reflects only ICRs whose supporting documents passed the keyword pre-filter, not the full corpus.

share a common real-world referent. The three regime-pairs (SORN–ICR, SORN–Inventory, ICR–Inventory) are treated as separate subtasks, because they differ substantially in difficulty: SORN–ICR pairs share regulatory context and tend to have stronger structural similarities, while ICR–Inventory pairs are often the most difficult because ICR filings describe data *collections* while the Inventory describes *systems*.

Following best practices (Jain et al. 2024; Beheshti et al. 2017; Zhao, Xue, and Min 2023), we do the following steps: (1) candidate retrieval; (2) probablistic scoring; (3) LLM reranking; (4) clustering, and (5) evaluation.

Candidate Retrieval The full cross-product of AI-relevant records (591 SORNs, 1,146 ICRs, and 4,558 AI Use Case Inventory entries) yields tens of millions of possible pairs across the three regime combinations; evaluating each individually is infeasible. Candidate retrieval reduces each pair space to a manageable set without discarding true links. For each record, the pipeline retrieves the 20 most similar records in the paired regime using an additive retrieval score:

$$s_{\text{retrieval}} = 0.5 \cdot s_{\text{TF-IDF}} + 0.2 \cdot s_{\text{agency\_name}} + 0.3 \cdot s_{\text{named\_entity}} \quad (1)$$

where tfidf is the cosine similarity between TF-IDF vectors (shared vocabulary per pair, up to 10,000 unigram and bigram features, log-scaled term frequency to reduce the influence of document length); agency is 1.0 if both records belong to the same parent agency and 0.0 otherwise; and ne is a normalized count of shared system acronyms, vendor names, and program names extracted via regular expression and named-entity recognition.

Two design choices follow from the structure of the data. First, whether or not the same agency filed the request is additive rather than a hard filter: cross-agency pairs with high TF-IDF similarity still appear in the top-20 candidates, because early results found cross-agency adoption cases, where a system developed by one agency is later used by another, among the most analytically important findings. Second, date proximity is not included in the retrieval score; qualitative review of early candidates found it to be a poor signal for links, since systems frequently generate regulatory filings years apart.

Probabilistic Scoring Each candidate pair surviving retrieval is scored on a weighted combination of features: whether the records come from the same agency, whether they share system names or acronyms, whether they mention the same vendors or programs, and how semantically similar their descriptions are. Formally:

$$s_{\text{composite}} = 0.25 \cdot s_{\text{agency}} + 0.25 \cdot s_{\text{name}} + 0.30 \cdot s_{\text{named\_entity}} + 0.15 \cdot s_{\text{semantic}} + 0.05 \cdot s_{\text{TF-IDF}} \quad (2)$$

where name is a fuzzy string similarity between system titles (exact match scores 1.0, shared acronyms score at least 0.8, otherwise token-set ratio); and semantic is the cosine similarity between OpenAI `text-embedding-3-small` embeddings. The TF-IDF cosine is retained at low weight (0.05) as a weak residual signal. Semantic similarity receives a lower weight than the structural features because regulatory language varies substantially across regimes and embedding models may not reliably capture domain-specific paraphrase relationships.

Pairs are assigned to one of three bands: scores above 0.62 advance to the LLM reranker as high-confidence candidates; scores between 0.30 and 0.62 advance as borderline candidates; and scores below 0.30 are rejected as confident non-links. Threshold values are calibrated against 8 known explicit SORN–ICR cross-references confirmed by document cross-citation.

LLM Reranking Pairs above the scoring threshold are submitted to a LLM for a classification with a structured explanation (see Extended version Appendix). The model is asked to classify each pair as same system, related (operationally connected but legally distinct—e.g., a data collection filing that feeds a separately described AI model), or no link, and to cite specific phrases from each document that support its decision. The "related" category matters analytically: it captures dependency relationships between systems that should not be collapsed into a single record but are still meaningful for understanding how disclosure gaps propagate across regimes.

Clustering Accepted same-system links from all three regime-pairs are aggregated into clusters using agglomerative clustering with group-average linkage, terminating when inter-cluster similarity falls below a threshold of 0.7

(Gao et al. 2024). This resolves a single AI system that appears across multiple regimes into one cluster rather than a set of disconnected pairs. A useful property of this approach is that cross-agency transitive links are recovered naturally: if a DHS system and a DOL system are each independently linked to a common AI Inventory entry, they are placed in the same cluster even if the DHS–DOL pair was never directly scored. Records classified as *related* by the reranker are not merged into clusters; they are preserved as a separate set of data-dependency edges.

Evaluation Because no labeled dataset exists for cross-regime AI system linking, we construct an evaluation set through annotation. We draw a stratified sample across three groups: links (cross-regime, same system), proposed related systems (shared data dependencies), non-links (pairs the system rejected with high confidence). One of the authors reviewed 100 random pairs of each of the groups and finds accuracy rates of: (1) 85% (same system); (2) 51% (related system); and (3) 100% (having no link).

5 System: GovAIfiles.org

We develop https://govAIfiles.org which contains a searchable web application and interactive data exploration features; each is described below. These components are available in a GitHub repository: https://github.com/triphora/govaifiles.org/.

Web application We build a web application that allows searching and filtering of AI Use Case Inventory records (see Figure 1). The application sets up a single database table with the consolidated AI Use Case Inventory records and associated metadata such as year published. The application imports the database table into our own search engine that only contains these records, then serves an API interfacing with our search engine. The search engine accepts two optional inputs: (1) a single query string, which will cause the output to only contain results with a match for the query in any text field of the AI Use Case Inventory records, and (2) a list of filters (e.g., year, agency, deployment stage), which will cause the output to only contain results matching every filter provided. The API interface directly passes the inputted query and filters to the search engine, and is publicly available for use. Detailed breakdowns of each AI Use Case Inventory are then easily parsable by individuals seeking information about government transparency documents (see Extended version Appendix).

Interactive data exploration. We include in the web interface a bar graph with interactive options to show the quantities and attributes of AI Use Case Inventory records. The graph shows the quantity of Inventory records reported by year and can be filtered and grouped by stage of development, high impact status, and agency/bureau.

Researcher backend. To enable deeper investigation of AI systems across disclosure regimes, we provide a cross-regime record linkage pipeline that links entries in the AI Use Case Inventory to related SORNs and ICRs. Using the method described in Section 4, we grouped items into cross-regime . The resulting link graph is available to researchers

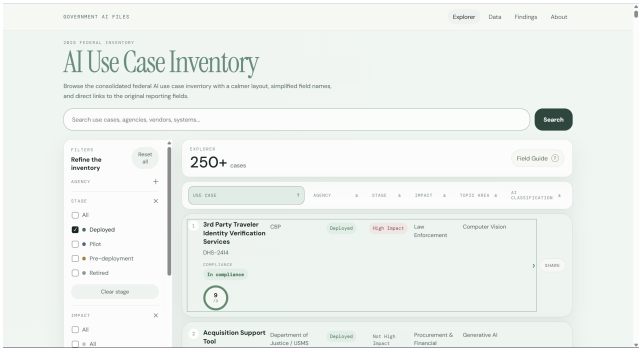

Figure 1: A screenshot of the home page of the govAIfiles.org search application.

as structured outputs that can be queried by system, agency, or regime pair. To illustrate the depth of analysis this enables, we construct an ATS (Automated Targeting System) subgraph (see Figure 2): seeded with AI inventory entries and one SORN for ATS, the pipeline surfaces almost 200 additional connected records across all three regimes, including ICRs that document data flows into ATS from immigration and customs forms.

6 Comparing Across Regimes: A Richer Picture of Federal AI Use

The developed cross-regime analysis methods allow us to create and identify clusters of interest. We focus on two examples that emerged from that analysis. The first, the Automated Targeting System (ATS), is a decision-support tool for border crossings characterized by not only some direct linkages across ICRs and SORNs, but also a rich graph of data dependencies where ATS feeds into another data collection or analysis system. The second, Mobile Fortify, the facial-recognition app, demonstrates that even when there are few direct textual matches, a broader cross-regime analysis can provide more context about data sources and governance.

Case Study: Automated Targeting System (ATS)

The Automated Targeting System (ATS) is CBP's decision-support tool for risk assessments for everyone who crosses a U.S. border. This case study highlights two findings. First, ATS exemplifies how bringing disclosures, fragmented across regimes, into conversation can reveal more information about government AI use. Second, ATS demonstrates the idea of mission creep, as it is a system, described in documentation with a narrow purpose, that has been repurposed and repackaged across uses and systems.

ATS emerged as one of the densest hubs in the clustering analysis, with connections spanning all three disclosure regimes (SORNs, ICRs, and AI Use Case Inventory) and a data-dependency structure that made it a natural candidate for closer examination (Figure 2 shows the subgraph). We define data dependencies as a link between systems where ATS is named as the processor or subject of data collection. The density of data dependency connections (dashed edges) relative to different parts of the same system (solid edges) is

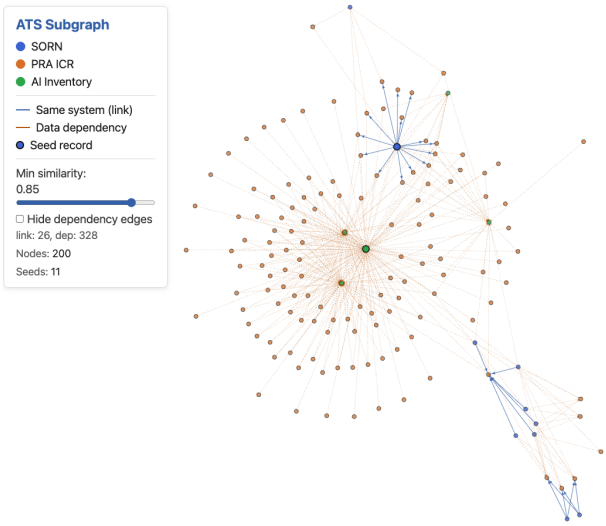

Figure 2: Nodes represent records from three federal disclosure regimes — SORNs (blue), ICRs (orange), and AI use case inventory entries (green). Solid edges connect records linked to the same underlying system across regimes; dashed edges indicate data dependencies, where one system draws from ATS. The graph reveals dense data-dependency structure relative to same-system links, alongside several tight clusters of SORNs and ICRs that document the same system under different disclosure obligations.

itself a finding: ATS is a frequent data source for AI-related systems across government.

ATS is a long-standing and controversial technology (Stanley 2007). Its governing Privacy Act notice (with SORN identifier DHS/CBP-006) was first published in 2006, describing a system that performed "rule based query of the cargo or personal information accessed by ATS" (Department of Homeland Security 2007). A 2012 update expanded the scope of data the system ingests, adding over dozen new source systems including the FBI's Terrorist Screening Database, but retained the same core vocabulary, describing ATS as "rule-managed technology that facilitates the targeting of high-risk travelers and cargo performing database comparisons, scenario-based targeting, and query functions" (Department of Homeland Security 2012). The notice has not been substantively updated since. Yet the 2025 AI Use Case Inventory discloses ten machine learning models operating within or feeding into ATS, none of which appear by name in the SORN. The SORN contains none of the words "artificial intelligence," "machine learning," or "algorithm." Whatever machine learning related capabilities ATS has acquired by 2025, the governing notice vocabulary has never been updated to reflect them.

Linking ATS to the AI Use Case Inventory allows information not provided elsewhere to be identified. Four of the ten models are classified as high-impact: the Passenger Security Assessment Model, Cargo Security Assessment Model, Illicit Trade model, and Traveler Entity Resolution. Two of the four, the Passenger Security Assessment Model and Traveler Entity Resolution, directly score individual travelers and draw on demographic variables including sex and age. All ten are categorized as Classical/Predictive Machine Learning or Computer Vision and designated as deployed systems. Their descriptions use language nowhere present in the Privacy Act notice: "advanced AI and machine learning models," "predictive modeling," "risk scores." Of the ten, two list ManTech as vendor (Passport Anomaly, DHS-81; Trade Entity Risk, DHS-95); the remaining eight are built in-house.

Beyond the AI Use Case Inventory, 166 ICRs, overwhelmingly from CBP and USCIS, name ATS as the downstream processor of collected data, tracing a supply chain that feeds models the SORN does not acknowledge. ATS deployments extend beyond CBP and USCIS. For example, two CDC records, governing COVID-era collection of airline passenger contact information, explicitly disclose that the data will be processed by "the DHS Automated Targeting System (ATS)." Several ICR supporting documents surface references to facial comparison technology operating within ATS, a capability the SORN never mentions (e.g., DHS/CBP/PIA-056, CBP/PIA-068). Moreover, a July 2017 ATS Privacy Impact Assessment (supplementing a 2022 ICR filing) partially fills in the post-SORN record, describing ATS as a modular decision-support system and acknowledging facial recognition and other AI components.

The cross-document accounting necessary to reconstruct who developed ATS technology and how it is used points to a structural problem in how AI transparency is currently framed. The conventional unit of analysis for AI disclosure is the model: a bounded artifact with an owner and a clearly delineated purpose. ATS resists that framing. The SORN describes a data storage and query platform; the inventory names the models; the ICR records trace the data flow from individuals. No single document connects those layers, and the 166 records naming ATS as a downstream processor suggest that the more analytically useful unit may be the data supply chain rather than the system itself. A traveler's information may flow through a form collection, into ATS, and be scored by a model that the SORN does not acknowledge, with no single disclosure document that follows that path end to end. That fragmentation is also what allows mission creep to go unaccounted for: a system first described in 2006 as performing "rule based query" of traveler information can by 2025 be running ten deployed machine learning models, four of them high-impact, processing data from agencies as far afield as CDC, without any major updated privacy or risk analyses that reflect the change.

Case Study: Mobile Fortify

Mobile Fortify is a DHS mobile application that captures facial images, fingerprints, and photographs of identity documents to verify the identity of individuals. It transmits those inputs to CBP, which runs the biometric matching; Mobile Fortify displays the returned results to ICE users. CBP owns the matching infrastructure; ICE operates the front-end mobile application. Mobile Fortify illustrates two limits of federal AI transparency infrastructure: (1) documentation lag that can appear as absence, and (2) gaps in disclosure such

that system-level understanding can remain incomplete even when cross-regime links are assembled.

The 2025 AI Use Case Inventory is the only transparency disclosure that names Mobile Fortify directly. It discloses two deployed, high-impact use cases: one for CBP and one for ICE. Both report that required pre-deployment safeguards (e.g., impact assessments, risk management protocols, and monitoring mechanisms) are still in progress. This potentially leaves it out of compliance with OMB's AI memorandum (Office of Management and Budget 2025).[2]

Mobile Fortify does not appear by name in either the SORN or ICR datasets. The two inventory records reference "TVSI" and "ATS-C" as associated systems, without definitions. "TVSI" most likely refers to TVS-Internal, the internal-facing variant of CBP's Traveler Verification Service. "ATS-C" is less obvious, plausibly a shorthand for ATS-Cargo, but no available document connects ATS-Cargo to Mobile Fortify's identity-verification workflow. What the ATS documentation does establish is that ATS and TVS are deeply integrated in CBP's biometric screening architecture, which Mobile Fortify appears to rely on. Cross-linking records helps, but it does not close the gap on its own.

Investigative reporting surfaced a February 2025 Privacy Threshold Analysis (PTA) that explains why no standalone Mobile Fortify SORN or PIA exists. CBP and ICE concluded that existing documentation provides sufficient coverage, despite acknowledging that Mobile Fortify is privacy-sensitive and requires both a PIA and a SORN (Cox 2025). The absence of a Mobile Fortify–specific notice is a deliberate determination, not an administrative oversight. We have no comparable reporting regarding the seeming lack of an ICR filing.

The disclosure record is both more available and less complete than it first appears. The AI Use Case Inventory names the application and describes two deployed, high-impact use cases; but its risk management fields are listed as in progress, and associated systems are identified only by opaque shorthand. Moreover, no information was publicly accessible during the months Mobile Fortify was operational before the January 2026 publishing of the Inventory. In addition to this time lag, we also identify that this deployed, high-impact system is operating while likely remaining non-compliant with its oversight requirements and vague in its details of integration with other government AI infrastructure.

7 Limits of Cross-Regulatory Comparison

Limited direct visibility of AI systems

Searching the corpus of 3,735 SORNs for terms drawn from our AI search configuration, including "artificial intelligence," "machine learning," "algorithm," and "automated decision" yielded few results. The term "artificial intelligence" appeared in only 6 SORNs; "machine learning" in 4. Broader proxy terms fared better: "algorithm" or "algorithmic" matched 34 SORNs, "risk assessment" matched 90, and "automated targeting" matched 30, for a total of 227 unique SORNs (6.1% of the corpus) containing at least one term from the search configuration. A zero-shot classification pass using a LLM identified 554 SORNs (14.8%) as AI-related when given access to full document text; a substantially higher figure., AI-related systems are documented in SORNs but without the vocabulary of AI disclosure.

Cross-referencing the 2023-2025 AI Use Case Inventories (710, 2,133 and 3,542 records, respectively) with the SORN corpus makes the divergence concrete: some AI inventory use cases describe systems that collect or process personal information and therefore should, under the Privacy Act, be covered by a SORN yet those SORNs describe the underlying record system without identifying it as AI-driven. 20 AI inventory records could be matched to a corresponding SORN; of those, none contain AI-related terminology.

The ICR corpus tells a similar same story. Our extraction pipeline covers 37,871 ICRs filed with OIRA from January 2018 through March 2026, with 171,560 supporting documents converted to searchable text. Even with this full-text infrastructure in place, a two-stage keyword filter; first identifying documents structured as OMB supporting statements, then filtering for any term from the AI search configuration, yielded only 1,160 candidate documents. LLM classification of the candidate set identified 1,144 ICRs as AI-related at high or medium confidence, representing a small percentage of the full corpus. As with SORNs, the dominant matched terms were functional descriptors: "risk assessment," "automated targeting," "algorithm", rather than the language of AI systems. The terms "artificial intelligence" and "machine learning" appeared rarely in supporting statements even for systems the AI Use Case Inventory explicitly labels as AI-driven, indicating that agencies filing ICR paperwork are not yet using the vocabulary of AI transparency, and that ICR disclosures will systematically undercount AI-adjacent systems under keyword-based searches.

Taken together, the low hit rates across all three corpora using explicit AI terminology reflect a structural feature of federal disclosure practice rather than an absence of AI systems: each regime has its own vocabulary, shaped by the legal mandate it serves. SORNs describe record systems under Privacy Act authority; ICR supporting statements justify information collection burden under the Paperwork Reduction Act; only the AI Use Case Inventory is organized around AI.The result is that even a researcher with full-text access to all three corpora, searching for AI by name, would substantially undercount the AI systems visible in the inventory.

Limited overlap across transparency regimes

Cross-regime analysis of explicit linkages reveals that government AI disclosures remain largely siloed, with minimal cross-references across the SORN, ICR, and AI Use Case Inventory frameworks. Of 554 AI-related SORNs, only 14 (2.9%) appear anywhere in the ICR corpus. Restricting the comparison to ICR filings that are themselves AI-related reduces that figure further, to 8 SORNs (1.6%). The remaining 97% of AI-related SORNs exist in isolation from any ICR filing. The AI Use Case Inventory fares no better: of all entries in the inventory, only two reference a SORN; both belonging to DHS's Automated Targeting System. After

[2]Due to the annual cadence of the disclosures, it's possible it became in compliance before OMB's deadline for agencies in April 2026 and after the 2025 Inventory was released in January 2026.

cross-document resolution, we find that between 80-97% of the documents (80% for SORNs, 97% for AI Use Case Inventory, and 82% for ICR records) contain no reference to another document (see Extended version Appendix figures).

The near-absence of cross-references does not, on its own, resolve whether the same systems appear across regimes. To recover implicit connections, we applied a multi-stage linking pipeline to the three AI-related corpora (554 SORNs, 1,146 ICR filings, and 5,268 AI inventory entries) to classify candidate pairs as the same underlying system, connected by a data flow or shared component, or unlinked. Even under this permissive standard, which does not require that agencies have acknowledged any connection, the rate of confirmed cross-regime matches is low: 446 pairs were classified as the same system across at least two regimes, representing 20% of AI-related records in the smaller of each paired corpus. A further 15,000 pairs were classified as sharing a data flow or common component—systems disclosed separately whose filings nonetheless describe a traceable dependency—a category that itself illustrates the fragmentation: a single capability may generate distinct records in each regime with no document pointing to the others.

The AI Use Case Inventory as an instructive tool

The federal AI Use Case Inventory is now in its third cycle. The 2023 inventory comprised 710 entries; by 2025 the program had grown to thousands of disclosures. Across the 2024 and 2025 inventories, 1,227 use cases appeared in 2024, 2,247 appeared for the first time in 2025, and 823 present in 2024 were absent from 2025 (see Extended version Appendix figures). Because neither inventory discloses persistent identifiers, year-over-year continuity was established via exact and fuzzy matching (see Extended version Appendix). The volume of new entries reflects continued expansion of federal AI activity. The 823 disappearances are harder to interpret; they may indicate the conclusion of operation of certain AI systems, reclassification, or something else entirely. Science and Administrative Functions account for the largest share of use cases, but high-impact designations are concentrated in Law Enforcement (which includes immigration enforcement) and Health & Medical (see Extended version Appendix figures). Of course, volume is an inadequate proxy for risk.

Among the 1,227 use cases traceable across both inventories, most deployed technology from 2024 remains active in 2025, but outcomes are not uniform (Figure 3). Pre-deployment systems most commonly remain in pre-deployment after a year. Among systems that did change stage, the most frequent transition is retirement. The third most common outcome is advancement: pre-deployment systems that are now deployed.

Considering the high-impact use cases and their required risk management disclosures, we find that a number of agencies and use cases fail to fill out the required fields. For example, there are 14 use cases that don't have an appeals process in place, and at least two use cases that have failed to respond to whether they have an AI impact assessment completed. In addition, qualitative review of the data shows that others have labeled required fields for high-impact deployed

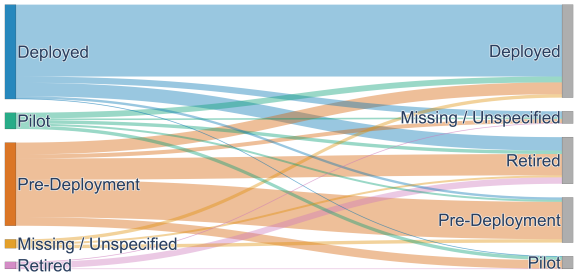

Figure 3: Sankey diagram demonstrating how the stages of development of the 1,227 AI use cases that were referenced in both the 2024 and 2025 Inventory changed. Most deployed AI technology continues to be in use in 2025, but heterogeneous outcomes occur.

use cases with "not applicable" without any additional justification and many still report that prerequisite impact assessment and risk management requirements are "in progress" rather than completed for already deployed systems.

8 Policy Recommendations

In considering how to make government transparency disclosures related to AI more effective and useful, we must also consider the audience and associated goals of such disclosures. Here, we are interested in disclosures of AI systems for a public audience with the goal of transparency into high-impact government operations. We thus focus our recommendations on improvements to the AI Use Case Inventory, as the existing mechanism that most closely matches this audience and goal. We note that, unfortunately, the AI use case inventory is currently set to no longer be required by statute as of 2027 (40 U.S.C. 2022), although the Executive Order on Trustworthy AI (Executive Office of the President 2020) may still be in force beyond that date. Thus, all these recommendations stem from our first key recommendation: **continue conducting the AI use case inventory in perpetuity.** While these recommendations are described in the context of this U.S. based transparency mechanism, they have broader implications for AI transparency.

Ensure clarity and broad applicability of the covered AI definition In the Inventories from 2023, 2024, and 2025 we see the results of three different variants of AI definitions and associated guidance. While all three years used the same core AI system definition, the 2024 inventory added clarifications to the definition (e.g., clarifying that it included small systems and basic machine learning tools like logistic regression) (Office of Management and Budget 2024). These clarifications were effective and drastically increased the number of systems reported (from 710 disclosed in 2023 to 2,133 in 2024). The 2025 inventory however appears to have included agency interpretations of the high-impact AI system definition that relied heavily on the part of the definition that limits such systems to those "whose output serves as a principal basis for a decision or action." The "principal

basis" language served to remove items from the inventory that under the 2024 interpretation were presumed included and appeared in the inventory. An AI use case can often serve as *a* basis for a decision or action, while questionably meeting the threshold of being a "principal basis." Broad applicability of such definitions is important for public insight into these tools, and for ensuring high-impact systems are covered by the required risk management steps; a simple fix in this case would remove the principal basis clause.

Create consistent AI system identifiers and cross-reference with other disclosure regimes The AI use case inventory is an annual disclosure regime; agencies are expected to report out on use cases each year, including those they've disclosed in a previous annual report. In such cases, it would help to the public's ability to track changes in use cases over time to have a common identifier per AI system that in each year's inventory. An identifier could also include the possibility of versioning via a main key for the AI system based on the use case and an associated version number, with directions given to agencies by OMB as to what constitutes a new version of the AI system. Such identifiers should also be referenced, as appropriate, in other disclosures, for example future SORNs could be required to list the AI use case identifiers for any AI systems expected to make use of the data. The creation of such identifiers is standard practice for other similar disclosure regimes, including SORNs and ICRs. Future AI inventories should also directly link to associated SORNs and ICRs. Creating links between these regimes would usefully allow the public a view into the full system that includes the AI system description, the data it relies on, and any previous privacy-related disclosures.

Provide public transparency into risk management processes In response to questions such as, "What are the key risks from using the AI for this particular use case and how were they identified?" the 2024 Inventory included some thoughtful descriptive examinations of potential risks and benefits of AI uses cases, although not all agencies released thorough assessments. Some by the Department of Veterans Affairs (VA) were notable for their nuanced and careful examination of risks and prevention of potential harms. For example, the VA created a risk assessment algorithm for opioid patients (Stratification Tool for Opioid Risk Mitigation, or STORM). In the 2024 inventory, key risks are detailed including an over-reliance by providers on the tool, and mitigations such as provider training are described; the full response is 367 words and identifies and responds to seven risks. Unfortunately, while the 2025 Inventory includes a question to agencies asking for potential impacts, no responses to that question were released publicly. This is a trend towards decreased length in responses across agencies visible in the number of words in the Inventory (see Figure 4). It's possible that agencies did still submit these more nuanced risk descriptions to OMB, but that OMB chose not to release those to the public this year. The public should have insight into these key risks and management processes.

Tailor inventory questions to the goal of transparency into high-impact use cases While transparency into risk management assessments and practices should be reinstated in the Inventory, these were only ever required for high-impact use cases. This appropriately puts agency reporting effort on cases already identified as high-impact. Other responses may be unnecessary for use cases that are not high-impact, or there may be ways to streamline agency responses so that they are both more useful and easier for the agency. For example, linking to SORNs and ICRs could help agencies fully describe underlying data used by a system without further paperwork. Other changes could be made via user testing the Inventory's questions with agencies themselves.

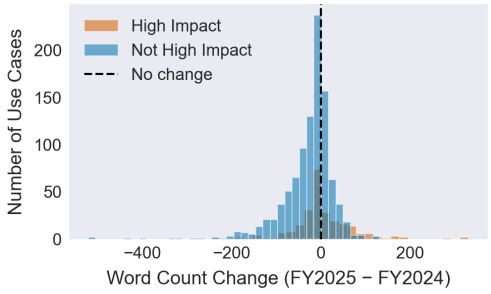

Figure 4: Overall we note a trend towards shorter responses in the AI use case entries that existed in both 2024 and 2025.

Make the inventory more publicly usable and archived In our view, a key goal of the AI use case inventory and the associated improvements we suggest here is public transparency. While the SORNs, ICRs, and AI use case inventory are technically public disclosures, they are not made available in a way that is usability tested and easy for the public to browse and interact with. SORNs and ICRs are released via the Federal Register, which is a fully text-based format and has the advantage of existing archival processes. The AI use case inventory is released on government websites, some of which have already disappeared in the changeover between the Biden and Trump Administrations, and on GitHub in CSV and Excel formats. Releasing the inventory with clearly comparable fields has allowed journalists following government technology changes to analyze the inventory and successfully provided some public transparency through their reporting (see, e.g., (Alder 2026)). However, the inventory should be maintained in an archival form and released via a website of the type we create (see Section 5) or other easily accessible format that allows searching by field or keyword, and which is user-tested with the general public.

9 Conclusion

Given that the AI Use Case Inventory is still relatively new, we should take advantage of the associated interpretive flexibility and continue to make improvements to make the disclosures more meaningful to the public. We describe how triangulating AI disclosures across different transparency regimes reveals more about AI systems despite structural limitations, and make policy recommendations to improve existing transparency regimes.

Acknowledgments

This work was supported in part by a grant from the Ford Foundation. This project also benefited from feedback from Chris Marcum, the Privacy Law Scholars Conference, University of Pennsylvania's Law & CS Roundtable, and the CS + Law Workshop.

References

40 U.S.C. 2022. Advancing American AI Act. Pub. L. No. 117-263, div. G, title LXXII, subtitle B, §§ 7224–7225, codified at 40 U.S.C. § 11301 note.

Alder, M. 2026. Disclosed government AI use increased by 70 *FedScoop*.

AlgorithmWatch and Bertelsmann Stiftung. 2020. Automating Society Report 2020. Technical report, AlgorithmWatch.

Ananny, M.; and Crawford, K. 2018. Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability. *New Media & Society*, 20(5): 973–989.

Bateyko, D.; and Levy, K. 2025. One Bad NOFO? AI Governance in Federal Grantmaking. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*, 1640–1652.

Beheshti, S.-M.-R.; Benatallah, B.; Venugopal, S.; Ryu, S. H.; Motahari-Nezhad, H. R.; and Wang, W. 2017. A systematic review and comparative analysis of cross-document coreference resolution methods and tools. *computing*, 99(4): 313–349.

Cox, J. 2025. You Can't Refuse To Be Scanned by ICE's Facial Recognition App, DHS Document Says. *404 Media*.

Cox, J.; and Cleek, A. 2025. How a US Citizen Was Scanned With ICE's Facial Recognition Tech.

Department of Homeland Security. 2007. Privacy Act of 1974; U.S. Customs and Border Protection, Automated Targeting System, System of Records. Notice of Privacy Act System of Records FR Doc. E7-15197, Department of Homeland Security, Office of the Secretary. Docket No. DHS-2007-0042. https://www.federalregister.gov/d/E7-15197.

Department of Homeland Security. 2012. Privacy Act of 1974; U.S. Customs and Border Protection, DHS/CBP-006-Automated Targeting System, System of Records. Notice of Privacy Act System of Records 77 FR 30297, Department of Homeland Security, Office of the Secretary. Docket No. DHS-2012-0019.

Executive Office of the President. 2020. Executive Order 13960: Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government. Federal Register. Document Number: 2020-27065.

Friedler, S. A.; and Selbst, A. D. 2025. The OMB Artificial Intelligence Memoranda. *Berkeley Technology Law Journal*.

Gao, Q.; Li, B.; Meng, Z.; Li, Y.; Zhou, J.; Li, F.; Teng, C.; and Ji, D. 2024. Enhancing cross-document event coreference resolution by discourse structure and semantic information. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 5907–5921.

Gellman, R. 2025. FAIR INFORMATION PRACTICES: A Basic History. https://ssrn.com/abstract=5348107.

Hartzog, W. 2017. The Inadequate, Invaluable Fair Information Practices. *Maryland Law Review*, 76(4): 952–982.

Hickok, M. 2024. Public Procurement of Artificial Intelligence Systems: New Risks and Future Proofing. *AI & Society*, 39(3): 1213–1227.

Jain, M.; Mutharaju, R.; Singh, K.; and Kavuluru, R. 2024. Knowledge-driven cross-document relation extraction. In *Findings of the Association for Computational Linguistics: ACL 2024*, 3787–3797.

Jobin, A.; Ienca, M.; and Vayena, E. 2019. The global landscape of AI ethics guidelines. *Nature machine intelligence*, 1(9): 389–399.

Kingsman, N.; Kazim, E.; Chaudhry, A.; Hilliard, A.; Koshiyama, A.; Polle, R.; Pavey, G.; and Mohammed, U. 2022. Public sector AI transparency standard: UK Government seeks to lead by example. *Discover Artificial Intelligence*, 2(1): 2.

Levendowski, A. 2021. Trademarks as surveillance transparency. *Berkeley Tech. LJ*, 36: 439.

Levy, A. L. 1994. The Paperwork Reduction Act of 1980: Unnecessary Burdens and Unrealized Efficiency. *Journal of Law and Commerce*, 14(1): 99–122.

Metcalf, J.; Moss, E.; Watkins, E. A.; Singh, R.; and Elish, M. C. 2021. Algorithmic impact assessments and accountability: The co-construction of impacts. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, 735–746.

OECD. 2019. Recommendation of the Council on Artificial Intelligence. Technical Report OECD/LEGAL/0449, Organisation for Economic Co-operation and Development. Adopted 22 May 2019; amended 3 May 2024.

Office of Management and Budget. 2024. Memorandum M-24-10: Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence. Memorandum M-24-10, Executive Office of the President, Washington, DC.

Office of Management and Budget. 2025. Memorandum M-25-21: Accelerating Federal Use of AI through Innovation, Governance, and Public Trust. Memorandum M-25-21, Executive Office of the President, Washington, DC. Rescinds and replaces M-24-10.

Pahlka, J. 2023. *Recoding America: why government is failing in the digital age and how we can do better*. Metropolitan Books.

Reisman, D.; Schultz, J.; Crawford, K.; and Whittaker, M. 2018. Algorithmic impact assessments: a practical Framework for Public Agency. *AI Now*, 9.

Ruggles, S.; and Magnuson, D. L. 2023. " It's None of Their Damn Business": Privacy and Disclosure Control in the US Census, 1790-2020. *Population and development review*, 49(3): 651–679.

Samaha, A. M. 2015. Death and Paperwork Reduction. *Duke Law Journal*, 65(2): 279–344.

Selbst, A. D. 2017. Disparate impact in big data policing. *Ga. L. Rev.*, 52: 109.

Selbst, A. D. 2021. An institutional view of algorithmic impact assessments. *Harv. JL & Tech.*, 35: 117.

Selbst, A. D.; and Barocas, S. 2018. The intuitive appeal of explainable machines. *Fordham L. Rev.*, 87: 1085.

Shapiro, S. 2012. The Paperwork Reduction Act: Research on Current Practices and Recommendations for Reform. Report, Administrative Conference of the United States, Washington, DC.

Shapiro, S. 2020. Reinvigorating the Paperwork Reduction Act. *Regulation*, 43(3): 36–41.

Stanley, J. 2007. ACLU Backgrounder on ATS. https://www.aclu.org/documents/aclu-backgrounder-ats.

Sunstein, C. R. 2013. The Office of Information and Regulatory Affairs: Myths and Realities. *Harvard Law Review*, 126(7): 1838–1876.

United States Congress. 1974. Privacy Act of 1974. Pub. L. 93-579, 88 Stat. 1896. 5 U.S.C. § 552a.

United States Office of Personnel Management. 2011. Paperwork Reduction Act (PRA) Guide. Technical report, United States Office of Personnel Management. Version 2.0.

U.S. Department of Health, Education, and Welfare. 1973. Records, Computers, and the Rights of Citizens: Report of the Secretary's Advisory Committee on Automated Personal Data Systems. Technical Report (OS) 73-94, Department of Health, Education, and Welfare, Washington, D.C.

Ustun, B.; Spangher, A.; and Liu, Y. 2019. Actionable recourse in linear classification. In *Proceedings of the conference on fairness, accountability, and transparency*, 10–19.

Wang, N.; McDonald, A.; Bateyko, D.; and Tucker, E. 2022. American Dragnet: Data-Driven Deportation in the 21st Century. Technical report, Center on Privacy & Technology at Georgetown Law.

Watkins, E. A.; Moss, E.; Metcalf, J.; Singh, R.; and Elish, M. C. 2021. Governing algorithmic systems with impact assessments: Six observations. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 1010–1022.

White House Office of Science and Technology Policy. 2022. Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People. Technical Report PREX 23.2:B 49, Executive Office of the President.

Zhao, J.; Xue, N.; and Min, B. 2023. Cross-document event coreference resolution: Instruct humans or instruct gpt? In *Proceedings of the 27th conference on computational natural language learning (CoNLL)*, 561–574.