Protecting Public Records from Automated Harvesting

AI, Automation, and the Threat to County Property Records

A Research Briefing for County Officials, Title Professionals, and Concerned Citizens

Prepared: March 2026

Executive Summary

Across the United States, county recorder offices, assessor databases, and court filing systems have moved online to increase government transparency and public access. This digitization, however, has created an unintended vulnerability: automated bots and AI-driven scraping tools can now harvest vast quantities of sensitive personal information—names, addresses, property ownership details, mortgage balances, legal descriptions, and more—at a scale and speed that was never anticipated when public records laws were written.

This briefing compiles current research, expert analysis, legislative developments, and real-world case studies to make a clear case: **county governments must act to restrict automated, bulk harvesting of public records**, particularly property and title documents, before widespread harm becomes irreversible.

This is not a question of restricting public access. Individual citizens, journalists, researchers, and professionals must retain their right to inspect government records. The question is whether bots and AI agents should be allowed to vacuum up entire county databases—containing the personal information of every property owner in a jurisdiction—without limitation, oversight, or accountability.

Critically, this problem has two sides. On one side, external actors are scraping county websites and exploiting public records systems. On the other—and this is the dimension most often overlooked—**counties themselves are enabling mass automated access** through bulk data sales programs, direct data feeds to commercial aggregators, open data portals with unrestricted API access, and vendor partnerships that place residents' information into commercial pipelines with no downstream oversight. Any effective response must address both the external threat and the internal compliance gap.

1. The Threat Landscape: What Is Happening Now

1.1 The Scale of Automated Scraping

The practice of web scraping—the automated extraction of data from websites—has evolved from a niche technical tool into an industrial-scale data harvesting operation. As professors Daniel J. Solove (George Washington University Law School) and Woodrow Hartzog (Boston University School of Law) argue in their landmark 2025 article in the California Law Review, the scale of scraping has reached a crisis point. Their article, awarded the Future of Privacy Forum's Privacy Papers for Policy Makers award, describes a fundamental clash between data harvesting and privacy principles.

> "Scraping violates nearly all of the key principles of privacy laws, including fairness, individual rights and control, transparency, consent, purpose specification and secondary use restrictions, data minimization, onward transfer, and data security." **— Daniel J. Solove & Woodrow Hartzog, "The Great Scrape," 113 California Law Review 1521 (2025)**

Their central argument is directly applicable to county public records: the fact that information is publicly available does not mean it is free for unlimited, automated extraction. As they write, scrapers treat all publicly available data as if it were "free for the taking," but privacy law has long recognized that public availability does not eliminate privacy interests.

1.2 County Records Under Siege: The Indiana Case Study

In January 2026, Indiana became one of the first states to directly address the problem of AI-driven harvesting of local government data through legislation. Testifying before the Indiana House government committee, John Wilson of the Allen County Board of Commissioners described how the county received commercial public records requests seeking bulk data exports of every purchase, vendor relationship, and financial transaction—formatted to the requester's own specifications.

> "This bill addresses a real and emerging threat that is happening largely out of public view. This threat is the large-scale harvesting of government data, often by bots or AI-driven systems outside of our state." **— John Wilson, Allen County Board of Commissioners, testimony before Indiana House Government Committee, January 21, 2026**

The proposed Indiana legislation would create a portal system allowing agencies to verify whether requesters are state residents or commercial entities, charge supplemental fees to out-of-state requesters, and prioritize requests filed for civic, journalistic, academic, or personal purposes over commercial bulk harvesting. As Wilson noted, many of the out-of-state requests came from fake names and companies that could not be verified.

1.3 Pennsylvania's Explosion of AI-Driven Records Requests

A March 2026 investigation by WPSU (Penn State's public media outlet) revealed that Pennsylvania is experiencing a dramatic surge in public records appeals—driven in significant part by AI. Liz Wagenseller, executive director of Pennsylvania's Office of Open Records, reported that appeals increased 64% in a six-month period compared to the same period the prior year, with AI playing a central role.

> "Around September, we started to see evidence of AI being used in appeals. If you compare the recent six months to six months a year ago at that same time, there's been a 64% increase in appeals—an enormous increase." **— Liz Wagenseller, Executive Director, Pennsylvania Office of Open Records (WPSU, March 2026)**

County governments bore the heaviest burden, with 705 appeals directed at county agencies in 2025—more than townships (535) or school districts (533). The County Commissioners Association of Pennsylvania described requests that are "overly broad, duplicative, or structured in ways that can delay or disrupt core government functions by overwhelming staff capacity."

2. The Other Side of the Problem: County-Enabled Automation

The public discussion of AI threats to public records tends to focus on external bad actors—scrapers, bots, and fraudsters attacking county systems from the outside. But there is an equally important and less discussed dimension: **counties themselves are actively enabling and facilitating automated, bulk access to the very data that puts their residents at risk.** This is not a case of the foxes breaking into the henhouse; in many instances, the door is being held open.

2.1 Bulk Data Sales Programs

Many county assessor and recorder offices operate formal bulk data sales programs. Riverside County, California, for example, maintains a public "Bulk

Data Sales” page on its Assessor-County Clerk-Recorder website, offering assessment records of real and personal property for purchase. While individual parcel lookups are available for free, the bulk data program allows commercial entities to acquire the county’s entire property database in a single transaction. This practice is common across the country and often occurs with minimal oversight of how the data will be used downstream.

2.2 Direct Data Feeds to Commercial Aggregators

The commercial property data industry is built on direct relationships with county offices. ATTOM Data Solutions, one of the largest property data aggregators in the U.S., openly advertises that it sources data from more than 3,000 U.S. counties and maintains a database of over 158 million properties—covering 99% of the nation’s population. Their database contains over 70 billion rows of transactional data and more than 9,000 discrete data attributes, including property tax records, deed information, mortgage details, foreclosure status, and ownership histories.

BatchData, another major aggregator, advertises “direct feeds from county assessor and recorder offices” that are updated daily. CoreLogic (now rebranded as Cotality) sources data from 99.9% of U.S. property records. These companies then resell this data through APIs, bulk licensing, and cloud platforms to anyone willing to pay—including marketers, hedge funds, insurance companies, and potentially bad actors who can use it for targeting vulnerable homeowners, phishing campaigns, or deed fraud reconnaissance.

The critical point for county officials is this: when a county provides direct data feeds to a commercial aggregator, it loses all visibility into and control over how that data is subsequently used. The data enters a commercial pipeline where it is blended with other sources, repackaged, and distributed through APIs that can be accessed by any paying customer worldwide. The county’s residents have no knowledge that their property information has been commercialized in this way and no ability to opt out.

2.3 Open Data Portals and API-First Policies

The “open data” movement has encouraged counties to publish datasets on public portals with full API access and bulk download capabilities. While the transparency goals of open data are laudable, the implementation often fails to distinguish between datasets that carry privacy risk and those that do not. Boulder County, Colorado, for example, licenses its open data—including parcel and property information—under a Creative Commons CC0 4.0 license, which

grants a worldwide, royalty-free, non-exclusive license to use, modify, and distribute the data for any purpose. Los Angeles County, San Diego County, San Mateo County, and Baltimore County all maintain similar portals.

Some municipal open data policies go further, explicitly requiring that datasets be published in formats that permit API access and bulk download. The City of Minneapolis's Open Data Policy, for instance, mandates that "all data sets published shall use a format that permits processing of the data for download through an automated programming interface (API) or bulk download." While this policy serves legitimate goals of government transparency, it treats all data uniformly—making no distinction between, say, public budget data and property records containing the home addresses and financial details of every resident.

2.4 eRecording Systems as Automation Infrastructure

The shift toward electronic recording (eRecording) of property documents has created another avenue for automated access. Companies like CSC maintain active eRecording relationships with over 3,600 counties nationwide. While eRecording offers genuine efficiency benefits for legitimate title and settlement professionals, the digital infrastructure it creates also means that property filings —which contain names, addresses, signatures, and legal descriptions—are entering fully digital pipelines from the moment of creation. Without proper access controls, this creates a situation where sensitive documents are "born digital" and immediately available for automated harvesting.

2.5 The Compliance Gap: What Counties Are Missing

The core problem is a compliance gap. Counties are modernizing their systems and data infrastructure—often under pressure from state mandates, vendor salespeople, or well-intentioned open data initiatives—without conducting adequate risk assessments of how automated access to their data could harm residents. Key questions that most counties are not asking include:

- Who is accessing our bulk data, and for what purpose?
- What contractual restrictions, if any, govern the downstream use of data we provide to commercial aggregators?
- Does our open data portal distinguish between low-risk datasets (budget reports, meeting minutes) and high-risk datasets (property ownership, addresses, financial details)?

- Are we tracking whether our data feeds are being used to build profiles that could facilitate fraud, stalking, or targeted marketing of vulnerable populations?
- Have we conducted a privacy impact assessment before signing data-sharing agreements with commercial vendors?

When counties enter into bulk data agreements or publish property records on open data portals without answering these questions, they are not merely failing to prevent a problem—they are actively contributing to it. The takeaway for county officials is not only “stop the scrapers” but also “audit your own data-sharing practices.”

2.6 The Scale of the Commercial Data Pipeline

To appreciate the magnitude of county-enabled data distribution, consider the numbers reported by just a few major aggregators:

- **ATTOM Data Solutions:** 158+ million properties from 3,000+ counties; 70+ billion rows of data; 9,000+ data attributes; delivery via API, bulk licensing, and cloud (Snowflake).
- **CoreLogic/Cotality:** Data from 99.9% of U.S. property records; 140+ million geocoded parcel maps; used by lenders, insurers, and government agencies.
- **BatchData:** 150+ million properties; direct daily feeds from county offices; delivery via API, S3, SFTP, and Snowflake.
- **Regrid:** Nationwide parcel data platform; sells parcel files by county and state; provides API and Esri-compatible services.
- **DataTree (First American):** 7+ billion searchable document images; 60+ search and filter attributes; used by lenders, title/escrow, and proptech companies.

Each of these companies built their databases by establishing relationships with county offices—relationships that counties entered into, often without fully understanding the downstream implications for resident privacy and safety.

3. Who Is Negatively Affected

3.1 Domestic Violence Survivors

Perhaps the most urgent public safety concern involves survivors of domestic violence and stalking. Property records contain home addresses—precisely the information that abusers seek to locate victims who have relocated for safety. The Electronic Privacy Information Center (EPIC) has documented this risk extensively.

> "For a domestic violence victim, the need for privacy is a need for physical safety. An abuser can take advantage of the general lack of protection for personal information in our society." **— Electronic Privacy Information Center (EPIC), Domestic Violence and Privacy**

The National Network to End Domestic Violence (NNEDV) has warned that public records—including court records and property filings—are increasingly published online, making it harder for survivors to maintain safety through relocation. When automated tools can scrape every property transfer in a county and cross-reference names across jurisdictions, a determined abuser can locate a survivor in minutes rather than months. Many states have address confidentiality programs, but these protections are rendered ineffective when bulk data harvesting makes the underlying records available outside government systems.

3.2 Elderly Homeowners

Seniors are among the most vulnerable targets of property fraud enabled by AI-driven data harvesting. The National Association of Realtors' 2025 Deed & Title Fraud Survey found that 63% of respondents nationwide were aware of title fraud occurring in their markets within the past 12 months. In the Northeast, that figure reached 92%. Fraudsters use AI tools to scan public records and identify vulnerable properties—particularly those owned by seniors, which are often mortgage-free and thus lack bank monitoring.

According to First American Title, a Florida title company encountered an AI-generated deepfake person during a video call meant to confirm the identity of a seller. The face used by the fraudsters belonged to a woman who had disappeared in 2018. Properties without an owner-occupant—vacant lots, second homes, rental properties—are the primary targets, but elderly owner-occupants with high equity are also at severe risk.

3.3 All Homeowners and Property Owners

The 2025 NAR survey found that 62% of known fraud cases involved vacant residential land, while 52% involved residential land specifically. Title fraud is more common in central cities (64%) and suburban areas (62%) compared to

smaller towns (40%). Fraudsters use forged documents—often generated by AI—to file transfers with county recorders. AI tools can generate convincing fake IDs, replicate notary stamps, and produce fraudulent deeds that fool county recording offices.

Four states (Arizona, Florida, Illinois, Utah) now require every county records office to provide electronic notification systems for property owners. Several counties in Missouri and Kansas have introduced “property title freeze” systems similar to credit freezes. Ten states empower county recorders to intervene when they suspect a fraudulent filing. But these defensive measures assume human-speed fraud—not AI-scale attacks.

3.4 County Government Staff and Taxpayers

Automated bulk requests impose enormous costs on county governments. Staff must locate, review, redact, and compile data across multiple departments. When bots generate hundreds of requests using fake identities, the cost falls on local taxpayers. The Pennsylvania County Commissioners Association reported that these requests place “significant demands on time and resources while ensuring sensitive information is properly protected.” Indiana’s proposed legislation directly addresses this by allowing supplemental fees for out-of-state commercial requesters.

3.5 Title Industry Professionals

The title industry—title examiners, abstractors, and closing agents—faces a dual threat. AI tools that scrape and analyze public records at scale threaten to displace workers who have traditionally performed this function, while simultaneously introducing new accuracy risks. Industry professionals note that AI struggles with the nuances of county recording systems: misspelled names in indices, grantors placed in grantee fields, and other human errors that experienced examiners catch routinely.

4. Risks and Public Safety Concerns

4.1 AI-Enabled Title Fraud at Industrial Scale

Deloitte’s Center for Financial Services projects that generative AI could drive U.S. fraud losses from \$12.3 billion in 2023 to \$40 billion by 2027—a compound annual growth rate of 32%. Real estate is a primary target. AI tools enable criminals to generate forged deeds, clone identities through deepfake technology,

craft targeted phishing emails impersonating county offices, and scan public records to identify vulnerable properties—all at unprecedented speed.

The American Land Title Association (ALTA) has documented how AI and deepfakes are reshaping fraud in real estate. Voice cloning, video deepfakes, and AI-crafted phishing emails can now create entire fabricated transaction chains that are nearly indistinguishable from legitimate business.

4.2 Mass Data Aggregation and the "Digital Biography"

Professor Solove has written extensively about how individually innocuous pieces of public information—when aggregated—create detailed personal profiles he calls "digital biographies." Property records contain: full legal names, home addresses, purchase prices, mortgage amounts, lender identities, legal descriptions of property, signatures, and notary information. When combined with other scraped public data (voter registration, court filings, business licenses), a complete financial and personal profile emerges.

> "The public availability of scraped data shouldn't give scrapers a free pass. Privacy law regularly protects publicly available data, and privacy principles are implicated even when personal data is accessible to others." **— Solove & Hartzog, "The Great Scrape," 113 California Law Review 1521 (2025)**

4.3 Infrastructure Burden and Denial of Service

Aggressive scraping can effectively constitute a denial-of-service attack on county websites. When bots consume bandwidth and server resources at industrial scale, legitimate users—citizens trying to look up a deed, title companies conducting searches, attorneys researching liens—are slowed or locked out. This undermines the very transparency that public records laws were designed to ensure.

4.4 Erosion of Public Trust

When citizens learn that their property records are being harvested by commercial entities and AI systems—often for purposes they never consented to—trust in government transparency erodes. The irony is acute: public records laws exist to keep government accountable to citizens, but the unintended consequence of digitization is that citizens' information becomes a commodity harvested by private actors for commercial gain or criminal exploitation.

5. Current Legislative and Regulatory Landscape

5.1 State-Level AI Governance (2025–2026)

The Center for Democracy and Technology (CDT) documented 50 public sector AI legislative proposals during the 2025 session alone, with activity focused on three areas: risk management practices, AI governance structures, and transparency requirements. Three standout laws passed in 2025:

- **Kentucky SB 4:** Established comprehensive public sector AI governance, including risk management policies for high-risk AI systems, an AI Governance Committee, mandatory public disclosure of AI use by agencies, and creation of an AI inventory.
- **Texas SB 1964:** Part of a slate of bills focused on public sector AI use, addressing procurement, oversight, and risk assessment for AI tools used by government agencies.
- **Montana HB 178:** Requires government agencies to disclose AI use, mandates human review for high-risk systems, and prohibits AI use by government for cognitive manipulation, social classification, deception, and surveillance in public spaces.

5.2 Indiana’s Anti-Bot Public Records Legislation (2026)

Indiana’s proposed legislation represents the most directly relevant model for county-level advocacy. Key provisions include: creating a portal system for public records requests that can distinguish bots from humans; requiring verification of a requester’s physical address; indicating to the public agency whether the requester is an Indiana resident; allowing supplemental fees for out-of-state commercial requesters (capped at 25 cents per page or \$25 per staff hour); and prioritizing requests filed by state residents for civic, journalistic, academic, or personal purposes.

5.3 Colorado Artificial Intelligence Act

Enacted in May 2024 with an effective date of June 30, 2026, Colorado’s AI Act is the most comprehensive AI law in the United States. It applies to both developers and deployers of AI systems doing business in Colorado and is designed to protect residents from algorithmic discrimination and misuse.

5.4 Title Fraud-Specific Legislation

Multiple states have enacted or are considering measures specifically targeting deed and title fraud:

- Four states (AZ, FL, IL, UT) require every county records office to provide property owner notification services.
- Five states (AR, CO, GA, IL, TX) require photo ID when filing real estate documents with the county.
- Ten states (CO, IL, MI, NV, NC, OH, OK, SC, SD, TX) empower county recorders to flag suspicious filings.
- Pennsylvania House Bill 1406 would create a specific criminal offense of deed fraud.
- Kansas and Missouri counties have introduced “property title freeze” systems allowing owners to lock their deed records.
- California has required Los Angeles County to report on the effectiveness of its notification system by 2028.

5.5 Federal Executive Action

In December 2025, President Trump signed Executive Order 14365, “Ensuring a National Policy Framework for Artificial Intelligence,” which establishes a policy to “sustain and enhance the United States’ global AI dominance through a minimally burdensome national policy framework.” The order mandates creation of an AI Litigation Task Force to challenge state AI laws deemed inconsistent with this policy. However, legal analysts note that executive orders cannot independently displace state legislation without congressional action, and the order specifically acknowledges that areas such as child safety protection and state government procurement will continue to be governed by state-level regulation.

5.6 The South Carolina Precedent

South Carolina’s Court Administration categorically banned automated scraping of court records. The ACLU of South Carolina and the state NAACP challenged the ban in federal court, arguing it unreasonably restricted First Amendment rights to access public information for research on eviction patterns. This case illustrates the tension at the heart of the issue: bans on scraping can restrict legitimate research, but unlimited scraping enables exploitation. The most effective policies

distinguish between human-initiated research and industrial-scale automated harvesting.

6. Expert and Professional Voices

6.1 Privacy and Law Scholars

Prof. Daniel J. Solove (George Washington University Law School) is one of the world's leading privacy scholars and author of over a dozen books on privacy law. His 2025 paper with Prof. Woodrow Hartzog, "The Great Scrape," received the Future of Privacy Forum's Privacy Papers for Policy Makers award. His earlier work, "Access and Aggregation: Public Records, Privacy, and the Constitution" (86 Minnesota Law Review, 2002), directly addresses how the aggregation of public records creates privacy harms that individual record access does not. He argues that government records containing personal information create "digital biographies" when assembled—and that existing public records frameworks were not designed for this aggregation.

Prof. Woodrow Hartzog (Boston University School of Law) specializes in privacy, media, and technology law. Together with Solove, he has proposed reconceptualizing the scraping of personal data as a form of surveillance—a framing directly applicable to the bulk harvesting of county property records.

6.2 Industry Professionals

Sarah Frano, Vice President and Real Estate Fraud Expert at First American Title, has documented how AI enables criminals to identify targets, scale their schemes, and avoid detection. She notes that deepfake technology can impersonate any party in a real estate transaction, including sellers, agents, and notaries.

Hany Farid, Professor at UC Berkeley and co-founder of GetReal Security, is considered a pioneer of digital image forensics. He helped develop PhotoDNA (used to fight child sexual abuse material) and has demonstrated real-time deepfake impersonation capabilities, warning that detection technology consistently lags behind generation technology.

6.3 Government and Regulatory Voices

Washington State has adopted Interim Guidelines for Purposeful and Responsible Use of Generative AI, acknowledging that "generative AI has the potential to catalyze innovation" but "must be deployed and regulated carefully to mitigate and guard against a new generation of risks, harms, and perpetuation of

existing inequities.” The guidelines recommend that local governments adopt their own policies around AI use.

7. The Case for County-Level Action

7.1 Why Counties, Specifically

County recorder offices are the custodians of property records. They are the point of entry for deed filings, the repository for title chains, and the public-facing database that both citizens and automated systems access. Counties have direct authority over how their digital systems are configured, what access controls are in place, and what terms of use govern their websites. This makes county government the most appropriate and effective level for intervention.

7.2 What “Banning Automation” Means (and Doesn’t Mean)

The goal is not to restrict public access to records. It is to ensure that access remains human-scaled rather than machine-scaled. Effective policies would:

1. Implement rate limiting and CAPTCHA systems on online record portals to prevent bulk automated harvesting.
2. Require identity verification for bulk or commercial data requests.
3. Distinguish between individual lookups and industrial-scale data extraction.
4. Prioritize requests from local residents and those with civic, journalistic, or personal purposes.
5. Charge cost-recovery fees for commercial bulk data access.
6. Mandate property owner notification systems for all deed-related filings.
7. Empower recorders to flag and hold suspicious filings for review.
8. Audit existing bulk data sales programs and commercial data-sharing agreements to assess downstream privacy risks.
9. Require privacy impact assessments before entering new vendor data-feed arrangements.
10. Classify datasets by risk level on open data portals, applying stricter access controls to records containing personal information such as addresses and financial details.

7.3 Bipartisan Appeal

This issue transcends partisan divisions. Property rights are a cornerstone of conservative policy priorities. Personal privacy and protection of vulnerable populations are priorities shared across the political spectrum. Fiscal responsibility demands that county resources not be consumed by bot-generated requests that serve commercial interests rather than civic transparency. Law enforcement and public safety concerns around fraud, identity theft, and stalking create natural allies in both parties. The Indiana legislation was introduced by Rep. Matt Lehman (R-Berne) with bipartisan support, demonstrating that this is an issue where both parties recognize the threat.

8. Key Statistics for Advocacy Materials

- **63%** of real estate professionals nationwide reported awareness of title fraud in their markets within the past 12 months (NAR 2025 Deed & Title Fraud Survey).
- **92%** of respondents in the Northeast reported awareness of title fraud (NAR 2025).
- **62%** of known fraud cases involved vacant residential land (NAR 2025).
- **64%** increase in public records appeals in Pennsylvania over a 6-month period, driven significantly by AI (PA Office of Open Records, 2026).
- **\$40 billion** in projected U.S. fraud losses enabled by generative AI by 2027 (Deloitte Center for Financial Services).
- **\$12.3 billion** in AI-enabled fraud losses in 2023, the baseline for Deloitte’s 32% annual growth projection.
- **83%** of surveyed real estate professionals said electronic notification systems for property owners are the most effective anti-fraud solution (NAR 2025).
- **705** public records appeals directed at county governments in Pennsylvania in 2025—more than any other type of government agency (PA Office of Open Records Annual Report).
- **50** public sector AI legislative proposals during the 2025 state legislative session alone (Center for Democracy and Technology).

- **3,000+** U.S. counties provide data to ATTOM Data Solutions alone, covering 158 million properties and 99% of the U.S. population.
- **70 billion+** rows of transactional property data and 9,000+ discrete attributes maintained by a single commercial aggregator (ATTOM), all sourced from county offices.
- **99.9%** of U.S. property records are available through CoreLogic/Cotality's commercial platform, sourced via relationships with county assessor and recorder offices.
- **150 million+** properties covered by BatchData through "direct feeds from county assessor and recorder offices" updated daily.

9. Source Index

Academic and Legal Scholarship

Solove, Daniel J. & Hartzog, Woodrow. "The Great Scrape: The Clash Between Scraping and Privacy." 113 California Law Review 1521 (2025). Available at scholarship.law.bu.edu and SSRN.

Solove, Daniel J. "Access and Aggregation: Public Records, Privacy, and the Constitution." 86 Minnesota Law Review 1137 (2002).

Industry and Professional Reports

National Association of Realtors. "2025 Deed & Title Fraud Survey." nar.realtor/research-and-statistics/research-reports/deed-and-title-fraud-survey

Deloitte Center for Financial Services. "Deepfake Banking and AI Fraud Risk." deloitte.com/us/en/insights/industry/financial-services/deepfake-banking-fraud-risk-on-the-rise.html

American Land Title Association. "AI, Deepfakes and the New Face of Fraud in Real Estate." alta.org (June 2025)

First American Title. "AI-Driven Fraud: The Hidden Threat in Real Estate." firstam.com

Experian Insights. "New AI Tools Facilitate Deed Fraud." experian.com/blogs/insights (November 2024)

Legislative and Government Sources

Indiana Capital Chronicle. “Indiana Public Records Access Law to Get Anti-Bot Overhaul.” indianacapitalchronicle.com (January 22, 2026)

Center for Democracy and Technology. “State Legislatures Continued Their Focus on Public Sector AI Use.” cdt.org (January 2026)

MOST Policy Initiative. “Real Estate Title Fraud.” mostpolicyinitiative.org (comprehensive state-by-state legislative analysis)

California SB 53 (2025–2026): Artificial Intelligence Models regulation. legiscan.com

Executive Order 14365 (December 11, 2025): “Ensuring a National Policy Framework for Artificial Intelligence.”

News and Journalism

WPSU (Penn State). “Fighting for Public Records Is More Common in Pennsylvania, Sometimes Aided by AI.” radio.wpsu.org (March 23, 2026)

Palo Alto Online. “How to Protect Your Property from Deed Fraud and Other Cybercrimes.” paloaltoonline.com (June 2025)

CalMatters. “California’s Report on AI Risks in Government Somehow Finds None.” calmatters.org (June 2025)

Advocacy and Civil Society

Electronic Privacy Information Center (EPIC). “Domestic Violence and Privacy.” archive.epic.org/privacy/dv/

National Network to End Domestic Violence, Safety Net Project. “Why Privacy and Confidentiality Matters for Victims.” techsafety.org/privacymatters

Reason Foundation. “State Bans on Website Data Scrapers Hinder Access to Public Information.” reason.com (June 2022; South Carolina ACLU/NAACP litigation)

Commercial Data Aggregators (County-Enabled Automation)

ATTOM Data Solutions. attomdata.com — 158M+ properties from 3,000+ counties; bulk data, API, and cloud delivery.

BatchData. batchdata.io — 150M+ properties; direct daily feeds from county assessor and recorder offices.

CoreLogic/Cotality. corelogic.com — 99.9% of U.S. property records; 140M+ geocoded parcel maps.

Regrid. regrid.com — Nationwide parcel data platform; county and state parcel file sales; Esri partnership.

DataTree (First American). datatree.com — 7B+ searchable document images from public land records.

Riverside County, CA Assessor-Clerk-Recorder. rivcoacr.org/bulk-data-sales — Example county bulk data sales program.

Boulder County, CO Open Data. bouldercounty.gov/government/open-data/ — Example CC0 4.0 licensed property data.

U.S. DOJ Final Rule: “Preventing Access to U.S. Sensitive Personal Data and Government-Related Data by Countries of Concern.” Federal Register, January 8, 2025. (Addresses national security risks of bulk data access including scraping of government data.)

End of Research Briefing

All sources cited in this document are publicly available.