COUNTY ACTION GUIDE

Protecting Public Records from Automated Data Harvesting

A Plain-Language Playbook for County Officials, IT Staff, and Recorders

Prepared: March 2026

No Technical Expertise Required — Start With Action 1 Today

How to Use This Guide

This guide is organized into three tiers based on how difficult each action is to implement. **You do not need to be a technology expert to use this guide.** Many of the most important steps are policy decisions, not technical ones. If your county has an IT department or a website vendor, they can handle the technical items—but you need to know what to ask them to do.

Tier 1 — Do This Week (No Technical Skills Required): Policy decisions, audits, and conversations you can start immediately. These are the items that cost nothing and require no technology changes.

Tier 2 — Do This Quarter (Basic IT Coordination Required): Changes to your website configuration and vendor relationships. Your IT staff or web vendor handles the implementation, but you direct what needs to happen.

Tier 3 — Do This Year (Strategic and Legislative): Longer-term actions including new policies, intergovernmental advocacy, and system redesigns. These require planning and budget, but they are the changes that create lasting protection.

> **Important:** Not every county will need every action. Use this guide as a menu, not a mandate. Start with Tier 1, then work through the items that apply to your county's specific situation.

Tier 1: Do This Week

These actions require no technical skills, no budget, and no new technology. They are policy audits and conversations that any county official can initiate immediately.

1.1 Audit Your Existing Data-Sharing Agreements

Most counties have agreements with commercial data companies—and many of those agreements were signed years ago without a full understanding of what happens to the data afterward. Your first step is to find out what you're already giving away.

What to do:

1. Ask your recorder, assessor, and clerk offices: "Do we sell or provide bulk data to any outside company?"
2. Collect copies of every data-sharing agreement, bulk data sales contract, and vendor partnership involving property records, assessor data, or deed filings.
3. For each agreement, answer these questions: Who receives our data? What are they allowed to do with it? Can they resell it? Is there an expiration date? Can we terminate it?
4. Make a list of every company that currently receives a direct data feed or bulk export from your county.

> **Why this matters:** Companies like ATTOM Data Solutions source data from over 3,000 U.S. counties. BatchData advertises "direct daily feeds from county assessor and recorder offices." Once your data enters these pipelines, it's repackaged and sold to anyone willing to pay—including people who may use it for fraud, phishing, or stalking. You need to know if your county is part of this pipeline.

1.2 Review Your Open Data Portal

If your county publishes data on a public open data portal, check whether property records with personal information are included—and under what license.

What to do:

1. Visit your county's open data portal (if one exists) and search for datasets containing property ownership, addresses, parcel data, or deed information.
2. Check the license terms. If it says "Creative Commons CC0" or "public domain," that means anyone in the world can download, redistribute, and commercialize that data with no restrictions.

3. Determine whether your portal distinguishes between low-risk data (budgets, meeting minutes) and high-risk data (property records, personal addresses).

4. If property records are published with unrestricted bulk download or API access, flag this for review.

> **Real example:** Boulder County, Colorado publishes parcel and property information under a Creative Commons CC0 4.0 license—a worldwide, royalty-free license for any purpose. This is common. If your county does the same, you are actively giving away resident data.

1.3 Check Your County Website's Terms of Use

Your county's website almost certainly has a Terms of Use page (sometimes called "Terms of Service" or "Acceptable Use Policy"). In many cases, this page says nothing about automated access, bots, or scraping. That's a problem, because without explicit terms, you have no legal basis to block a scraper even if you detect one.

What to do:

1. Find your county website's Terms of Use page. If there isn't one, that's your first gap.

2. Check whether it explicitly prohibits automated access, bot scraping, bulk data extraction, or use of the website by non-human agents.

3. If it doesn't, draft language (see Section 5 of this guide for sample language) and begin the process to adopt it.

> **Quick win:** Even a basic Terms of Use update—prohibiting automated scraping and bulk harvesting—gives your county a legal tool it currently lacks. Courts have recognized Terms of Use violations as a factor in scraping litigation.

1.4 Talk to Your Recorder About Fraud Alerts

Four states (Arizona, Florida, Illinois, Utah) already require every county recorder to offer property owner notification services. Even if your state doesn't mandate it, many county recorders can set up a basic alert system. The NAR's 2025 survey found that 83% of real estate professionals consider notification systems the single most effective anti-fraud measure.

What to do:

1. Ask your recorder: “Do we offer a free property owner alert system? If someone files a document against a resident’s property, does the owner get notified?”
2. If the answer is no, ask what it would take to implement one. Many vendors offer this as an add-on to existing recording systems.
3. If the answer is yes, ask how it’s promoted. Many counties have these systems but almost no residents know about them.

1.5 Identify Your County’s Most Exposed Data

Not all public records carry the same risk. A county budget spreadsheet poses no privacy threat. A database of every property owner’s name, home address, purchase price, and mortgage balance is a gold mine for fraudsters. You need to classify your data by risk level.

What to do:

1. List every dataset your county makes available online, through portals, through bulk sales, or through vendor feeds.
2. For each dataset, ask: Does it contain home addresses? Full names? Financial information? Signatures? Legal descriptions tied to individuals?
3. Mark datasets as High Risk (contains personal/financial data), Medium Risk (contains names or addresses without financial detail), or Low Risk (no personal information).
4. High Risk datasets should be the first priority for access controls in Tier 2.

Tier 2: Do This Quarter

These actions require coordination with your IT department or website vendor. You direct the policy; they handle the implementation.

2.1 Add a robots.txt File to Your County Website

What it is: A robots.txt file is a simple text file that sits on your website and tells automated bots which parts of the site they are allowed to visit. Think of it as a “No Trespassing” sign for robots. It won’t stop determined bad actors (just as a

sign won’t stop a determined trespasser), but it blocks all well-behaved bots—including the ones operated by major AI companies and search engines—from vacuuming up your data.

What to tell your IT staff or web vendor:

1. We need a robots.txt file in the root directory of our county website.
2. It should block known AI training bots (GPTBot, ClaudeBot, Google-Extended, CCBot, anthropic-ai, Bytespider, and others) from accessing our property records, assessor data, and recorder document search pages.
3. It should allow regular search engines (Googlebot, Bingbot) to index our general public-facing pages so citizens can still find us through Google.
4. Provide them the specific directory paths where property records, document searches, and assessor lookups are hosted so they can write targeted Disallow rules.

> **Limitation:** robots.txt is a request, not a command. Well-behaved bots (Google, Bing, major AI companies) respect it. Malicious scrapers will ignore it. It’s the equivalent of locking the front door—it won’t stop a burglar, but it stops everyone who would have walked in without thinking.

2.2 Implement Rate Limiting on Record Search Pages

What it is: Rate limiting means setting a cap on how many searches or page views a single visitor can perform in a given time period. For example: “No more than 20 property lookups per minute from the same computer.” A normal citizen doing a single lookup will never notice. A bot trying to download your entire database will be stopped cold.

What to tell your IT staff or web vendor:

1. We need rate limiting on our property record search pages, assessor lookup tools, and document image viewers.
2. A reasonable starting point: 20–30 requests per minute per IP address for search queries; 5–10 document downloads per minute per IP address.
3. When a visitor exceeds the limit, they should see a friendly message explaining the limit and suggesting they contact the office for bulk data needs.

4. The system should log IP addresses that repeatedly hit the limit so we can identify persistent scrapers.

> **Benefit:** Rate limiting is invisible to normal users and costs nothing to implement on most web platforms. It immediately stops industrial-scale scraping while preserving normal public access.

2.3 Add CAPTCHA to Record Search Pages

What it is: A CAPTCHA is the “I’m not a robot” checkbox or image puzzle that many websites use. It forces a human to interact with the page before accessing records. Bots cannot solve CAPTCHAs (or can only solve them at significant cost and delay).

What to tell your IT staff or web vendor:

1. We want a CAPTCHA challenge on our online property record search, deed image viewer, and assessor data lookup pages.
2. Use a modern, accessibility-compliant CAPTCHA (Google reCAPTCHA v3, Cloudflare Turnstile, or hCaptcha). Avoid old-style distorted-text CAPTCHAs—they frustrate users and are also less effective against modern bots.
3. The CAPTCHA should appear before search results are displayed, not after. This prevents bots from reading the data and simply discarding the CAPTCHA page.
4. Ensure the CAPTCHA meets ADA accessibility standards—audio alternatives must be available for visually impaired users.

2.4 Move Sensitive Records Behind Authentication

What it is: For the highest-risk records (deed images, mortgage documents, documents containing signatures), consider requiring users to create a free account before accessing them. This is not gatekeeping—the records remain free and public—but it creates a login wall that bots cannot easily bypass and creates an audit trail of who accessed what.

What to tell your IT staff or web vendor:

1. We want to require free account registration to view or download deed document images and recorded instrument images.

2. The registration should require a valid email address and basic identity verification (name, reason for access).
3. We still want assessor data (property values, basic ownership info) to remain accessible without an account—the account requirement is specifically for document images containing signatures, notary stamps, and detailed financial instruments.
4. All access should be logged so we can audit usage patterns.

> **Balance transparency with safety:** This approach preserves full public access—anyone can create a free account in two minutes—while eliminating anonymous bulk harvesting. It follows the same model used by PACER (federal court records) and many state court systems.

2.5 Renegotiate or Terminate Risky Vendor Agreements

Based on the audit you conducted in Tier 1, you now know which companies receive your data. This quarter, renegotiate or terminate agreements that don't include adequate protections.

What to require in any data-sharing agreement going forward:

- The vendor must disclose all downstream recipients of your county's data.
- The vendor may not resell or redistribute individual-level data (names + addresses) without your county's written consent.
- The agreement must include a termination clause allowing your county to revoke access if the vendor's practices change.
- The vendor must notify your county within 30 days of any data breach involving your county's records.
- The agreement must specify permitted uses and explicitly prohibit use for targeted marketing to vulnerable populations, phishing, or fraud-related activities.

2.6 Remove or Restrict Bulk Download Capabilities

If your county's online portal currently allows anyone to download the entire property database in a single click (via a "Bulk Download" button, CSV export, or API endpoint), this should be restricted immediately.

What to do:

1. Remove any “Download All” or “Export Full Database” buttons from public-facing pages.
2. If an API exists, require an API key (a unique identifier assigned to approved users) for access. This creates an audit trail and allows you to revoke access if the key is misused.
3. For legitimate bulk data needs (title companies, licensed professionals), create a formal request process that requires identification and a stated purpose.

Tier 3: Do This Year

These are strategic, longer-term actions that require planning, budget, and collaboration with other agencies and elected officials.

3.1 Adopt a Formal Data Governance Policy

Your county needs a written policy—approved by your board of commissioners or supervisors—that governs how public records data is stored, shared, sold, and protected. Without a written policy, every decision is ad hoc, and protections disappear when staff turns over.

The policy should address:

- Data classification: Which datasets are High Risk, Medium Risk, and Low Risk (building on the audit from Tier 1).
- Access tiers: What level of access is available to the general public, to licensed professionals, and to commercial entities—and what controls apply at each tier.
- Vendor oversight: Requirements for any company receiving county data, including downstream use restrictions, breach notification, and audit rights.
- Open data standards: Which datasets belong on an open data portal and which require gated access.
- Regular review: A schedule (at least annual) for reviewing all data-sharing agreements and access controls.

- Incident response: What happens when a scraping attack or data misuse is detected.

3.2 Implement a Tiered Access System

The most effective long-term solution is a tiered access system that provides different levels of access based on who is requesting and why. This is not about restricting public access—it is about distinguishing between a citizen looking up one deed and a bot trying to download every deed in the county.

The tiers might look like this:

- **Public Tier (no account required):** Basic property information—owner name, address, assessed value, tax status. Available to anyone, with rate limiting and CAPTCHA in place.
- **Registered Tier (free account required):** Full deed images, recorded documents, mortgage details. Requires a free account with email verification. Rate-limited to prevent bulk downloads.
- **Professional Tier (verified identity + stated purpose):** Higher-volume access for title companies, attorneys, real estate professionals, and journalists. Requires professional license verification or a signed acceptable use agreement.
- **Commercial Tier (formal agreement + fees):** Bulk data access for commercial entities. Requires a signed data-use agreement with downstream restrictions, cost-recovery fees, and audit provisions. Modeled on Indiana’s proposed legislation.

3.3 Lobby Your State Legislators

County-level action is essential, but state legislation creates a uniform standard that protects all counties at once. Indiana’s 2026 anti-bot legislation provides a ready-made template. Your county should actively lobby for similar legislation in your state.

Key provisions to advocate for:

- A portal system that can distinguish automated requests from human ones.
- Verification of requester identity (physical address, residency status).

- Supplemental fees for out-of-state commercial requesters (Indiana proposed 25 cents per page or \$25 per staff hour).
- Priority processing for requests filed by state residents for civic, journalistic, academic, or personal purposes.
- Authority for county recorders to flag and hold suspicious filings.
- Mandatory property owner notification systems statewide.

> **Bipartisan framing:** Present this as a property rights and fiscal responsibility issue. Conservative officials respond to property rights protection and reducing taxpayer burden from bot-generated requests. Progressive officials respond to protecting domestic violence survivors and preventing predatory targeting of vulnerable populations. Both care about stopping fraud.

3.4 Join or Form an Intergovernmental Working Group

No single county can solve this alone. Aggregators work across thousands of counties—they will simply route around any one county's restrictions. The most powerful action is coordinated action.

What to do:

- Contact your state's County Commissioners Association, County Recorders Association, or equivalent body and propose a working group on automated data harvesting.
- Share your audit findings (Tier 1) with neighboring counties—they likely have the same vendors and the same gaps.
- Develop model policies and model contract language that any county in your state can adopt.
- Coordinate with the National Association of Counties (NACo) to elevate this issue to the national level.

3.5 Conduct a Privacy Impact Assessment

Before entering any new technology contract, vendor agreement, or open data initiative, your county should conduct a formal Privacy Impact Assessment (PIA). A PIA is a structured process for identifying what personal data is involved, who will access it, what the risks are, and what safeguards are needed.

A PIA should answer:

- What personal data is being collected, stored, or shared?
- Who will have access to this data, and through what channels?
- What is the potential harm if this data is misused? (Consider domestic violence survivors, elderly homeowners, and identity theft victims.)
- What safeguards are in place to prevent misuse?
- Is there a less invasive way to accomplish the same goal?

3.6 Invest in Staff Training

Technology solutions only work if your staff understands why they matter. Many county employees process records requests on autopilot—they may not recognize when a request is coming from a bot, when a bulk request is suspiciously formatted, or when a filing contains signs of AI-generated fraud.

Training should cover:

- How to recognize bot-generated records requests (identical formatting, fake names, out-of-state addresses, requests for data in unusual formats).
- How to recognize AI-generated fraudulent documents (inconsistencies in notary stamps, unusual font rendering in signatures, metadata anomalies).
- When and how to flag a suspicious filing for review.
- Your county’s new data governance policy and what it means for daily operations.
- Why these protections matter—real stories of domestic violence survivors located through property records, elderly homeowners defrauded through AI-generated deeds, and counties overwhelmed by bot-generated requests.

4. How These Actions Protect Jobs

A common concern is that restricting automation will slow down legitimate work. The opposite is true. Uncontrolled automation is the threat to jobs—not the restrictions on it.

4.1 The Threat to Title Industry Professionals

Title examiners, abstractors, and closing agents have spent careers developing expertise in county recording systems. They know the quirks—misspelled names in indices, grantors placed in grantee fields, jurisdictional variations. AI tools that scrape and analyze records at scale threaten to displace these workers, but they do so with less accuracy and no accountability. By restricting bulk automated access, counties preserve the market for human expertise while still allowing professionals to do their work through normal channels.

4.2 The Threat to County Staff

When bots generate hundreds of public records requests using fake identities, county staff must process each one—locating, reviewing, redacting, and compiling data across departments. This consumes staff time that should be spent serving real residents. Pennsylvania's Office of Open Records reported a 64% increase in appeals driven by AI, with county governments bearing the heaviest burden (705 appeals in 2025). Restricting automated requests doesn't eliminate jobs—it protects existing staff from being overwhelmed by machine-generated workload.

4.3 The Competitive Advantage of Human Judgment

When counties implement tiered access systems, they create a market advantage for licensed professionals who work through proper channels. A title company that accesses records through an authenticated professional tier—with human review—can offer something that a bulk data scraper cannot: accuracy, accountability, and local knowledge. This is good for consumers, good for the industry, and good for county revenue.

5. Sample Language and Templates

5.1 Sample Terms of Use Language (Website)

Add the following to your county website's Terms of Use page:

"Automated access to this website—including but not limited to web scraping, crawling, data mining, or the use of bots, spiders, or other automated tools to extract data—is prohibited without the express written consent of [County Name]. This prohibition applies to all pages, databases, document images, and search functions hosted on this website. Violation of this policy may result in permanent blocking of the offending IP address, referral to law enforcement, and civil action for damages. This restriction does not apply to standard search engine indexing of general public-facing pages."

5.2 Sample Data-Sharing Agreement Clause

Include the following in any agreement with a commercial data recipient:

“Recipient shall not redistribute, resell, sublicense, or otherwise transfer individual-level data (defined as data that identifies or could be used to identify a specific natural person, including but not limited to name, home address, financial details, or property ownership information) to any third party without the prior written consent of [County Name]. Recipient shall maintain an audit log of all access to county-provided data and shall make such log available to [County Name] upon request. Recipient shall notify [County Name] within thirty (30) days of any breach, unauthorized access, or misuse involving county-provided data. This agreement may be terminated by [County Name] with sixty (60) days’ written notice for any reason.”

5.3 Sample Board Resolution Language

Use the following as a starting point for a board resolution directing staff to implement protections:

“WHEREAS, [County Name] maintains public records containing the personal information of its residents, including home addresses, property ownership details, financial information, and signatures; and WHEREAS, automated bots and AI-driven systems are capable of harvesting this information at a scale and speed that threatens resident privacy, enables fraud, and overwhelms county resources; and WHEREAS, the [County Board/Commission] has a responsibility to protect residents while maintaining transparency; NOW THEREFORE BE IT RESOLVED that the [County Administrator/IT Director/Recorder] is directed to: (1) conduct an audit of all existing data-sharing agreements involving property records; (2) implement technical safeguards including rate limiting and CAPTCHA on online record search systems; (3) update the county website’s Terms of Use to prohibit automated scraping; and (4) report back to this Board within 90 days with findings and recommendations.”

6. Implementation Checklist

Use this checklist to track your county’s progress. Check off each item as completed.

Tier 1 — This Week

☐ Audit all existing data-sharing and bulk data sales agreements

☐ Review open data portal for high-risk property datasets

☐ Check website Terms of Use for anti-scraping language

☐ Confirm whether recorder offers property owner fraud alerts

☐ Classify all publicly available datasets by risk level

Tier 2 — This Quarter

☐ Direct IT/vendor to add robots.txt blocking AI training bots

☐ Implement rate limiting on property record search pages

☐ Add CAPTCHA to online record search and document viewer

☐ Move deed document images behind free account registration

☐ Renegotiate or terminate risky vendor agreements

☐ Remove or restrict bulk download and unrestricted API access

Tier 3 — This Year

☐ Adopt formal data governance policy (board-approved)

☐ Design and implement tiered access system

☐ Engage state legislators on anti-bot records legislation

☐ Join or form intergovernmental working group

☐ Conduct privacy impact assessment for all data systems

☐ Implement staff training program on AI fraud and bot detection

---

End of County Action Guide

For questions or assistance implementing these actions, visit publicrecordssafety.com.