FRAME + FORGE
AI STRATEGY STUDIO

HOW AI TOOLS QUIETLY EXPOSE SRH DATA

A plain-language guide for the reproductive health movement: how a single visit becomes data that gets sold — even when nobody breaks a rule. Built for SRH leadership in a post-Dobbs, post-*Purl* landscape. Read once. Share widely. Act this week.

By **Lyndsay Sanborn, MHPA** · Frame + Forge AI Strategy Studio · April 2026

CONTENTS · WHAT'S IN THIS HANDOUT

01 · WHAT HIPAA ACTUALLY IS — AND WHAT IT ISN'T

HIPAA only protects when two things are true

HIPAA is the federal law that protects what it calls **Protected Health Information (PHI)**. People hear "HIPAA" and think the data is locked in a vault. It isn't. HIPAA only kicks in when both of these conditions apply at the same time:

CONDITION 1

The data identifies a specific patient

Name, date of birth, address, phone, medical record number — the kind of details that point to one person. **Strip those off, and HIPAA stops protecting it.**

CONDITION 2

It's held by a covered entity or business associate

A "covered entity" (clinic, hospital, health plan) or its "business associate" — a vendor working under a Business Associate Agreement (BAA). **Hand it to anyone outside that circle, and HIPAA stops protecting it too.**

The whole game lives here:

Modern AI features make both of these conditions stop being true on a regular basis — and most clinics never see it happen. The rest of this handout shows you exactly **how**.

02 · THE TWO KINDS OF AI INSIDE YOUR EHR · DIFFERENT WORK, DIFFERENT RISK

GENERATIVE AI · WRITES THINGS

It writes, summarizes, transcribes, and drafts.

Drafts the chart note from a recorded visit. Writes after-visit instructions. Transcribes the patient-provider conversation in real time.

EVERYDAY EXAMPLE

Like ChatGPT — it puts words on a page based on patterns it learned.

MAIN PRIVACY RISK

The contract usually lets the vendor use your patients' visits to train the next version of the AI — which then gets sold to other clinics.

PREDICTIVE AI · CREATES INFERRED DATA

It assigns probability scores about the patient.

"70% likely pregnant." "Elevated substance-use risk." "Possible intimate partner violence." Privacy professionals call this **inferred data** — new information the AI creates about the patient that the patient never said.

EVERYDAY EXAMPLE

Like how Netflix predicts which movies you'll like — except about your pregnancy, mental health, or substance use, and the prediction is stored forever.

MAIN PRIVACY RISK

03 · THE BUSINESS MODEL YOU DIDN'T KNOW ABOUT

Many AI vendors can make money twice from your patients' data.

When your clinic signs up for an AI feature, you become a customer in two ways at once. One you can see on the invoice. The other you cannot. Both are usually written into the same standard contract you signed.

WAY 1

Subscription Revenue

Your clinic pays the vendor a monthly or annual fee for the AI feature. **This is the money you know about.**

WAY 2

Selling Your Data

Vendors can strip names off the data flowing through the AI and **sell it to data brokers, pharma companies, insurers, and analytics firms.** This revenue stream is invisible to you.

STEP 1
Your Clinic
Patient visits happen, AI runs.

▶

STEP 2
The Vendor
Strips names. Now "de-identified."

▶

STEP 3
Data Broker
Combines with other datasets.

▶

STEP 4
The Buyer
Pays for the recombined data.

WHO BUYS THIS KIND OF DATA, AND WHY

- **Pharma companies** — for drug marketing & clinical research
- **Health insurers** — for risk scoring and underwriting
- **Advertisers** — to target ads at patients by health status
- **Other data brokers** — to enrich their own datasets and resell
- **Population health analytics** — resold to all of the above
- **State actors / law enforcement** — via subpoena to the broker, not the clinic

WAIT — ISN'T "DE-IDENTIFIED" THE SAME AS ANONYMOUS?

No. "De-identified" means "legal to sell." It does not mean unidentifiable.

HIPAA's **Safe Harbor** rule says: remove 18 specific **direct identifiers** (name, address, dates more specific than year, phone, email, MRN, and 13 others) and the data is no longer protected. A second pathway, called a **"Limited Data Set,"** lets vendors keep dates and 3-digit ZIP under a Data Use Agreement. Either way, the **indirect identifiers** stay behind — ZIP, visit timing, demographic patterns, encounter type. Data brokers already have your phone's location data, your purchase history, your address from public records. **Combining indirect identifiers + auxiliary data lets buyers trace it back to a real, named person** — and they don't break any law to do it.

PICTURE THIS

The vendor uses a Limited Data Set (legal under HIPAA). The row they sell reads: **female, age 28, ZIP 942xx, visit type: pregnancy test, date: April 12, 2026.** A data broker already has the location data of every phone that pinged a tower near that clinic that day. A small clinic might see four patients in that window. Cross-reference against any consumer dataset and you're back to a name. **This is how re-identification actually works in 2026.**

This isn't hypothetical — it's already happening.

- **FTC v. GoodRx (2023):** $1.5M penalty for sharing prescription & health data with Facebook and Google for ad targeting.
- **FTC v. BetterHelp (2023):** $7.8M penalty for sharing mental-health intake data with advertisers.
- **Texas AG v. Seattle Children's (2023):** AG issued a Civil Investigative Demand for records on out-of-state patients seeking gender-affirming care.
- **FTC v. Premom (2023):** Period-tracking app shared user pregnancy and fertility data with marketing companies.
- **Meta & Nebraska (2022):** Facebook handed over private messages used to prosecute someone for self-managed abortion.
- **Major AI scribe vendor contracts** (Abridge, Nabla, Suki, Microsoft DAX, Augmedix) often include model-training or de-identification clauses; terms vary, and some vendors are tightening defaults.

~ 1 in 3

SHADOW AI IN CLINICAL PRACTICE

U.S. physicians reported using AI in clinical practice in 2024 — up sharply from prior years. **A meaningful share** use it without their organization's knowledge or a Business Associate Agreement. *(Sources: Doximity 2024 State of AI in Medicine; AMA digital health surveys; figures approximate.)*

HIPAA compliance and patient privacy are not the same thing.

THE LINE WE WANT EVERY LEADER TO REMEMBER

BEFORE THE MAPS · FOUR TERMS TO HOLD

The vocabulary you need to read the journeys below.

Inferred Data
New information the AI **creates** about a patient (pregnancy probability, IPV risk, substance-use score) — not facts the patient stated.

Safe Harbor · 18 IDs
HIPAA's rule for stripping 18 **direct identifiers** (name, address, dates more specific than year, phone, MRN…). After stripping, data is "de-identified" and legal to sell.

Indirect Identifiers
ZIP, visit timing, demographics, encounter patterns. They **stay behind** after Safe Harbor stripping — and they're enough to re-identify when combined with broker data.

The Gray Zone
Where inferred data lives **once it leaves the chart**: protected inside the EHR, often unprotected once it flows to vendor analytics.

04 · PATHWAY A · INSIDE THE EHR

Follow Jordan through one routine visit

Meet Jordan.

They're 28 and book a routine visit at their local SRH clinic. The clinic runs a popular EHR — think *Epic, athenahealth, Oracle Cerner, eClinicalWorks, or NextGen* — with an AI scribe layered on top, like *Abridge, Nabla, Suki, Microsoft DAX, or Augmedix*. All widely deployed. The clinic signed a standard BAA. Staff are trained. A privacy officer is on payroll. Everyone follows every rule. Here's exactly what happens to Jordan's data — and why each step is a problem.

01

THE VISIT

JORDAN SEES
A normal conversation. They may not know an AI scribe is on.

BEHIND SCENES
Audio captured, transcribed, fed to prediction models.

WHY IT MATTERS
Jordan never consented to AI use or knew where the data would go.

02

INFERRED DATA

JORDAN SEES
Nothing. The AI works silently.

BEHIND SCENES
Predictive features generate risk scores — pregnancy, IPV, substance use. This is **inferred data**.

WHY IT MATTERS
This is new data Jordan never said. It now exists in the gray zone.

03

THE SPLIT

JORDAN SEES
A new note in the portal. Same as always.

BEHIND SCENES
Note → EHR (protected). Inferred data → vendor analytics.

WHY IT MATTERS
The most sensitive data just entered the gray zone.

04

DATA SOLD

JORDAN SEES
Their life continues. They are never notified.

BEHIND SCENES
Vendor strips direct IDs (Safe Harbor) or uses a Limited Data Set. Either way, legal to sell.

WHY IT MATTERS
Indirect identifiers (ZIP, visit timing, demographics) stay — enough to re-identify.

05

RE-COMBINED

JORDAN SEES
Maternity ads on their phone. They never searched for them.

BEHIND SCENES
Broker combines with location, purchase, and cellphone data. Jordan is identifiable again.

WHY IT MATTERS
"De-identified" was a legal label, not real anonymity.

Every step was legal. Jordan is still exposed.

Then a state attorney general subpoenas the data broker — not the clinic. HIPAA doesn't apply to brokers, so they comply. **The clinic is never told. Jordan is never told.** The whole exposure happened because the inferred data (information the AI created about Jordan that Jordan never said) traveled outside the chart, became "de-identified," got sold, and got reassembled downstream. **This is not a hypothetical. This is the architecture.**

Now follow Dr. Lee, who never opened the EHR

Meet Dr. Lee.

The clinician who saw Jordan that morning. After Jordan left, Dr. Lee opened ChatGPT on their phone to look up something specific about Jordan's case — age, pregnancy status, presenting symptoms. The clinic doesn't know. There's no Business Associate Agreement with OpenAI. Here's exactly what happens to that one query — and why.

01

THE QUESTION

DR. LEE SEES

A helpful chatbot explaining Jordan's case.

BEHIND SCENES

Jordan's details just went to a third party.

WHY IT MATTERS

The clinic never agreed to share data with this company.

02

NO AGREEMENT

DR. LEE SEES

A clean chat window. Nothing unusual.

BEHIND SCENES

Free ChatGPT, Claude, Gemini: **no BAA exists**.

WHY IT MATTERS

Without a BAA, sharing PHI with the vendor is prohibited.

03

LIKELY BREACH

DR. LEE SEES

A useful answer in seconds.

BEHIND SCENES

Age + pregnancy + symptoms = identifiable. **In most readings, a reportable breach.**

WHY IT MATTERS

Narrow gray areas exist — the clinic owes a documented risk assessment either way.

04

LOGGED FOREVER

DR. LEE SEES

They close the tab and forget.

BEHIND SCENES

Without a BAA, every clinician query is logged on vendor servers in fully readable text. May be used for training, depending on vendor and tier.

WHY IT MATTERS

A hostile state AG can subpoena these logs — vendor isn't required to notify clinic, clinician, or patient.

05

INSIDE THE MODEL

DR. LEE SEES

Nothing — this is invisible to them.

BEHIND SCENES

Jordan's details may become part of the trained model.

WHY IT MATTERS

Months later, fragments could surface to another user.

Worse than Pathway A. There's no agreement at all.

The biggest risk isn't the training data — it's the legal exposure. ChatGPT logs every clinician query in fully readable text. Without a BAA, those records can be subpoenaed by a hostile state attorney general — and the vendor isn't required to notify your clinic, your clinician, or your patient. Under HIPAA's Breach Notification Rule, this is, in most readings, also a reportable breach (narrow gray areas exist; the clinic still owes a documented risk assessment). **This is happening right now in clinics that don't know it's happening.**

06 · THE LEGAL FLOOR JUST GOT LOWER

The 2024 federal protections for reproductive health data are gone.

In **April 2024**, HHS finalized a HIPAA rule giving reproductive health information extra protection from law-enforcement disclosure. In **June 2025**, a federal court in *Purl v. HHS* vacated that rule nationwide. Baseline HIPAA still applies. State laws still apply where they exist. **The extra federal layer is gone — and unlikely to return soon.**

What this means in practical terms: the contracts you sign with AI vendors and the governance work your organization does are doing more of the protective work now than they were a year ago. The law is no longer the safety net.

A SPECIFIC DANGER AFTER *PURL V. HHS*

ChatGPT logs every query — and a hostile state AG can subpoena them.

Every prompt typed into a consumer AI tool (ChatGPT, Claude, Gemini, Copilot) is stored on the vendor's servers, often indefinitely, and is fully readable text — not encrypted from the vendor. **Those records can be obtained by law enforcement with a subpoena or search warrant.**

A state attorney general investigating out-of-state travel for reproductive or gender-affirming care can request OpenAI, Anthropic, or Google's logs the same way they request data broker records. **The vendor is not required to notify your organization, your clinician, or your patient.** The first you may hear about it is in a court filing — or never.

07 · THE STATE LAW LANDSCAPE

Where you operate matters more than it did a year ago

With the federal floor gone, state law is doing the protective work. It varies wildly. A patient who travels across state lines effectively passes through different privacy regimes.

STRONGEST · WASHINGTON

My Health My Data Act

The toughest health privacy law in the country. **Covers inferred health data.** Requires consent for sale or sharing of consumer health data. Private right of action.

STRONG · CALIFORNIA

CMIA + CCPA + AB 254

CMIA covers health data; AB 254 specifically extends protection to **reproductive and sexual health data**. Restricts cooperation with out-of-state investigations.

WEAKEST · TEXAS

HIPAA baseline only

No state-level health privacy law beyond HIPAA. **Active state AG pursuing data on out-of-state travel** for reproductive and gender-affirming care via subpoenas to brokers.

08 · SIX QUESTIONS LEADERSHIP SHOULD BE ASKING

Plain-English questions any executive can ask

1. **What inferred data is this AI generating about our patients that they never said out loud?**
   Predictive scores are new data the patient didn't give you.
2. **Where does that data go after the visit, and who can see it?**
   Map every system the chart, the scores, and the telemetry touch.
3. **Is our patient data being used to train the vendor's AI?**
   If yes, your visits are teaching a product that gets sold to others.
4. **What is the vendor allowed to do with the de-identified version?**
   "Any lawful purpose" means it can be sold and combined with broker data.
5. **If a subpoena comes to the vendor, will they tell us before they hand anything over?**
   Notification has to be contractual now — the federal layer is gone.
6. **Do our clinicians use ChatGPT or OpenEvidence without a BAA?**
   If yes — or if you don't know — that's where to start.

09 · WHAT TO DO THIS WEEK

Six concrete steps any leadership team can take now

1. **Run an AI inventory**
   List every tool in your org that touches patient data — including the ones IT didn't procure. You can't govern what you can't see.
2. **Read your BAAs**
   Find the de-identification, model-training, and product-improvement clauses. Those are the levers. Most leaders have never read them.
3. **Survey your clinical staff**
   Ask them, anonymously, which AI tools they use for clinical work. The answer will surprise you. The gap is your shadow-AI risk.
4. **Issue a one-page staff guidance**
   Tell staff what AI use is OK and what is not yet. Give them a sanctioned alternative. Bans alone don't work — workflows do.
5. **Schedule a board conversation**
   AI governance is now a board-level risk. Make the AI inventory + vendor accountability a standing agenda item, not a one-time review.
6. **Plan for state-law variance**
   If you serve patients who travel, your governance has to assume a Texas-baseline floor. Build the contracts and policies for the worst case.

10 · IF YOU NEED HELP

Frame + Forge does this work specifically for the reproductive health movement.

AI inventories. Vendor accountability reviews. Board-ready governance frameworks. Built by someone with 25+ years in SRH, not a generic compliance shop. **The first conversation is free.**

TALK TO LYNDSAY →

The bottom line. Compliance protects you from regulators. Governance protects your patients. You need both — and the technology is moving faster than the law. *This handout uses composite scenarios. Every mechanism in it is real and currently in use.*