TECHNOLOGY · DATA PROTECTION

Protect Patient, Trial, and Research Data in AI Tools

An OmniTek Consulting implementation guide for VPs and Directors in pharma, biotech, and medical device companies.

READING TIME
About 7 minutes

WHO IT'S FOR
VPs and Directors

GOES WITH
The AI Maturity Quick Check

> The fastest way to lose leadership support for AI is a data incident: patient details, trial data, or an unpublished result pasted into the wrong tool. A few clear rules, a short checklist for staff, and a plan for when something goes wrong prevent most of it.

SCOPE This guide covers AI used in how your company works, such as drafting, analysis, and operations. It doesn't cover AI that is part of a product you sell, such as AI-enabled device software or software as a medical device. That has its own regulatory path; check it with your Regulatory team.

Why it matters

- **The data is regulated and contracted.** Patient and trial participant data comes with privacy law, consent limits, and partner contracts attached.
- **Research data is the business.** Unpublished results, formulations, and process know-how are often a company's most valuable assets.
- **Mistakes are easy.** It takes one copy and paste to move a case narrative or a draft clinical study report into a tool that isn't approved for it.

What counts as sensitive

Type	Data class	Examples	Default rule
Patient and health information	Personal and patient	Adverse event narratives, patient support records, medical information requests with identifiers	Only tools approved for this class; remove identifiers wherever possible

Type	Data class	Examples	Default rule
Trial participant data	Personal and patient	Case report form data, coded subject data, eligibility details	Only within approved systems and the limits of consent and contracts
Unpublished research	Confidential	Study results, draft clinical study reports, assay data, formulations	Approved enterprise tools with no training on your data, team-limited access. Extra rule: results under a partner or licensing confidentiality clause, and topline results not yet publicly disclosed, go only into named tools, or not into AI tools at all. Participant-level data is Personal and patient.
Partner and licensor information	Confidential	Data and documents shared under a confidentiality agreement	Check the agreement before any AI use
Employee personal data	Personal and patient	Performance notes, HR cases	Keep out of general AI tools

Coded (pseudonymized) trial data may still be personal data under GDPR, depending on who can link it back to the participants. The EU Court of Justice has said pseudonymized data is not personal data in all cases and for every person (Case C-413/23 P; Court press release). Treat coded data as personal data unless your privacy team concludes otherwise.

Controls that work

Control	What it means
Approve tools by data class	Every approved tool has a written list of the data classes it may receive
Minimum necessary	Enter only what the task needs; summarize or excerpt rather than paste whole records
De-identify first	Remove names, dates, and other identifiers before using AI on patient or case content, using your company's de-identification standard
Check consent and contracts	Before using trial or partner data in a new way, confirm the consent and agreements allow it
Company accounts only	Sensitive work only in company-managed accounts with sign-on and logging, never personal ones
Technical guardrails	Where available, use data loss prevention and admin settings that block or flag sensitive content

A before-you-paste checklist

Give this to everyone who uses AI tools. It fits on a sticky note.

TEMPLATE · COPY AND ADAPT

Before you paste anything into an AI tool: [ ] Is this tool approved for this kind of data? [ ] Have I removed names, initials, dates of birth, and subject or case numbers I don't need? [ ] Is any of this unpublished research or a partner's confidential information? [ ] Am I in my company account, not a personal one? [ ] Will I check the output before I use it? If any answer is no or unsure: stop and ask [AI owner or Privacy contact].

Removing these is basic minimization, not HIPAA de-identification. The HIPAA Safe Harbor method requires removing 18 types of identifiers, or you can use Expert Determination instead (HHS de-identification guidance).

If something goes wrong

If sensitive data goes into a tool it shouldn't have:

1. **Stop and don't delete anything yet.** Note the tool, the time, and what was entered.
2. **Report it right away** to your Privacy or data protection contact and the AI owner, using your company's incident process.
3. **Let Privacy and IT decide next steps**, such as asking the vendor to delete the data and assessing whether notification is required. Your privacy and legal team decides whether notification applies, and deadlines can be short. Under the GDPR, the supervisory authority is notified without undue delay and, where feasible, within 72 hours of becoming aware of a breach, unless it is unlikely to result in a risk (Article 33; Commission summary). Under HIPAA, covered entities notify affected individuals without unreasonable delay and no later than 60 days after discovery, and business associates notify the covered entity no later than 60 days after discovery (HHS).
4. **Fix the cause.** Update the checklist, training, or tool settings so it doesn't happen again.

Make reporting easy and blame-free. People who are afraid to report make incidents worse.

Your first 30 days

1. **Week 1: Agree the data rules with Privacy and Quality.** Confirm which data types count as sensitive and which approved tools may receive each.
2. **Week 2: Check your tools.** Make sure each approved tool's data classes are written down, and turn on any available guardrails.
3. **Week 3: Brief the team.** Walk through the checklist with real examples from your department, and hand it out.

4. **Week 4: Test the incident path.** Walk through a made-up incident with Privacy and IT, so everyone knows who does what.

By day 30 your team knows what's sensitive, which tools can take it, and what to do if something goes wrong. That covers the data protection question in the Quick Check at the Operationalizing level; enforcing the rules through system controls moves you toward Scaling.

Common mistakes

- **Relying on memory.** A policy no one can recall in the moment doesn't help. The checklist is the reminder.
- **Treating coded data as anonymous.** Coded or pseudonymized trial data is often still personal data.
- **Forgetting partner agreements.** Confidentiality terms can rule out AI use even when your own policy allows it.
- **Punishing reporting.** Teams that fear blame hide incidents. Make it easy to speak up.
- **Approving a tool without naming its data classes.** "Approved" alone invites people to paste anything.

Regulatory touchpoints

Which rules apply depends on your company's role, markets, and contracts. Status is as of October 2026.

Area	What it means for AI use
HIPAA (US), including its de-identification standards	Applies where your department handles protected health information as a covered entity or business associate. A vendor needs a business associate agreement only when it is a business associate: it creates, receives, maintains, or transmits protected health information on behalf of a HIPAA covered entity (HHS)
GDPR (EU) and similar privacy laws	Personal data, which can include pseudonymized trial data, needs a lawful basis and a data processing agreement with the vendor. Health data, including trial participants' health data, is a special category and also needs a GDPR Article 9 condition
Informed consent and trial agreements	Secondary use of trial data is limited by the informed consent, the agreements between sponsors and sites, and applicable privacy law; check with Privacy before reusing it
Confidentiality agreements	Partner and licensor data may carry restrictions on processing or third-party tools

This guide is general information, not legal or regulatory advice. Your Privacy, Legal, and Quality leads should confirm what applies to your data.

How OmniTek can help

The full AI Maturity Report Card measures security and data protection alongside the other maturity dimensions, and the **vendor evaluation kit** includes the privacy and contract questions to ask before approving a tool.

To talk through protecting your data, or your Quick Check results, book time with OmniTek.

Book time with OmniTek

This guide is general information for leaders in regulated life sciences functions, not legal or regulatory advice. Confirm requirements with your Quality, Regulatory, and Legal teams. © 2026 OmniTek Consulting.