AIMONGER
WHITE PAPER

AI Discovery

The Privacy Gate: AI for Law Firms Under GDPR, the EU AI Act, and Professional Secrecy

The constraint is real but it is not a ban. The line that matters is controller versus processor: consumer AI products make the vendor a controller of your client's data with no DPA and possible training use, while commercial tiers reframe the vendor as processor with a DPA and no training. A local PII gate adds a second defence by pseudonymising prompts before they leave your boundary. Malta's Professional Secrecy Act makes the stakes higher than for a generic EU firm because client confidentiality applies even when the data is not personal data under GDPR.

David Saliba . 2026-08-28 . aimonger.com/whitepapers/ai-privacy-law-firm-gdpr-gate/

TL;DR

Yes, you have a real legal constraint. No, it is not "AI is banned for law firms." The constraint is that the **consumer** versions of Claude, ChatGPT, Copilot, DeepSeek, and GLM make the vendor a data controller of your client's data (no Data Processing Agreement, possible model training, retention up to five years), which collides with both the GDPR (Regulation (EU) 2016/679) and Malta's Professional Secrecy Act (Chapter 377 of the Laws of Malta). The **commercial and enterprise** tiers reframe the vendor as a processor with a DPA and no training, which handles most of the GDPR mechanics contractually, but you still carry Maltese professional secrecy, DPIA (Data Protection Impact Assessment, GDPR Article 35), and human-oversight obligations on top. A **local PII gate** (a local model that pseudonymises prompts before they reach any cloud API) adds a structural defence that no contract clause can match.

Good: The constraint is real and your instinct is correct.

Bad: The framing "just running Claude or Copilot" is too blunt. The exposure depends entirely on which tier of which product you use and where the data goes.

Because: GDPR (the EU's General Data Protection Regulation, Regulation (EU) 2016/679) draws a hard line between controller (the party that determines the purposes and means of processing) and processor (the party that processes on the controller's instructions). The vendor's role flips based on product tier. Malta's Professional Secrecy Act adds a second, stricter layer that GDPR alone does not capture, because it protects client confidential information even when that information is not "personal data" in the GDPR sense.

What I did: Researched current (2026) privacy documentation and terms of service for Anthropic (Claude), OpenAI, Microsoft (Copilot), DeepSeek, and Zhipu AI (GLM). Cross-referenced the Ganado Advocates Chambers AI 2026 Malta guide for the Malta-specific legal framework. Pulled primary court records for Bartz v. Anthropic and the Italian Garante's enforcement orders against OpenAI and DeepSeek. Reviewed open-source local PII gate tools (Hey Jude, Septum, pii-proxy, localscrub) for the mitigation architecture.

Glossary of terms

Term	Definition
Controller	The natural or legal person who determines the purposes and means of processing personal data (GDPR Art. 4(7)). The law firm is always a controller for client data.
Processor	A natural or legal person who processes personal data on behalf of the controller (GDPR Art. 4(8)). The AI vendor is a processor only under commercial tiers with a DPA.
DPA	Data Processing Agreement (also called Data Processing Addendum). The Article 28(3) contract between controller and processor that sets processing instructions, confidentiality, security, and sub-processor rules.
DPIA	Data Protection Impact Assessment (GDPR Art. 35). A structured risk assessment required before high-risk processing, including large-scale processing of personal data through AI.
GDPR	General Data Protection Regulation, Regulation (EU) 2016/679. The EU's primary data protection law, directly applicable in Malta.
SCCs	Standard Contractual Clauses. EU-approved contract terms that provide a legal basis for international personal data transfers to third countries (GDPR Chapter V).
ZDR	Zero Data Retention. A vendor configuration where inputs and outputs are not stored after the API response is returned. Available from Anthropic and OpenAI for qualifying enterprise customers.
EUDB	EU Data Boundary. Microsoft's commitment to process and store EU customer data within the EU.
PII	Personally Identifiable Information. Information that can identify a natural person, alone or in combination.
Pseudonymisation	Processing personal data so it can no longer be attributed to a specific data subject without additional information kept separately (GDPR Art. 4(5)).
Professional secrecy	The legal duty of lawyers to keep client information confidential, codified in Malta's Professional Secrecy Act (Chapter 377). Broader than GDPR because it covers non-personal client information.
Shadow AI	Employee use of consumer AI tools (ChatGPT, Claude free, DeepSeek) on client data without the firm's knowledge or approval. The most common compliance failure.
PIPL	Personal Information Protection Law. China's data protection law (2021). Subordinate to state security interests, unlike GDPR.
EDPB	European Data Protection Board. The EU body that issues binding guidance on GDPR interpretation across member states.
IDPC	Information and Data Protection Commissioner (Malta). Malta's data protection authority.
MDIA	Malta Digital Innovation Authority. Malta's lead AI regulator under the EU AI Act implementation.
LLM	Large Language Model. The underlying AI system behind tools like Claude, ChatGPT, and GLM.
NER	Named Entity Recognition. A technique for detecting names, dates, locations, and other entities in text. Used in PII gate detection.
Open-weight model	A model whose trained weights are published under a permissive licence (e.g. MIT), allowing self-hosting on your own hardware.

The problem in one sentence

If a law firm cannot say, for each AI tool its lawyers use, whether the vendor is a controller or a processor of the client data that flows through it, the firm is already non-compliant and does not know it.

Why this briefing exists

Lawyers are being asked to adopt AI for efficiency while keeping client confidentiality, legal professional privilege (LPP, the right protecting lawyer-client communications), and professional secrecy intact. The market default is pasting into the consumer chat box that is already open, and that path is non-compliant. The compliant path is choosing the right product tier, signing a DPA, running a DPIA, and installing a local PII gate that strips personal data before it leaves the firm's boundary.

This dossier is written for the managing partner who must approve the AI policy, the DPO (Data Protection Officer) who must evidence it, and the IT lead who must implement it. It covers five major AI providers, the full EU and Malta legal framework, the court cases and data breaches that prove the risk is not theoretical, the runaway decisions that cascade into an audit, and the local-model gate architecture that makes compliant AI practical.

The legal framework

2.1 EU-level: GDPR (Regulation (EU) 2016/679)

The GDPR is the load-bearing statute. It applies directly in Malta as an EU member state and is supplemented by the Data Protection Act (Chapter 586 of the Laws of Malta). The articles that matter for AI use in a law firm:

Controller and processor (Articles 4(7), 4(8), 28). The law firm is the controller. The AI vendor is the processor only if a DPA is in place that meets the Article 28(3) requirements (documented instructions, confidentiality obligations, security measures, sub-processor rules, assistance with data subject rights). Without a DPA, the vendor is not acting as your processor - it is acting as an independent controller of whatever data your lawyers paste in.

Lawful basis (Article 6). The firm needs a lawful basis for processing client personal data through an AI tool. In practice this is typically client consent (Article 6(1)(a)) or legitimate interests (Article 6(1)(f)) subject to a balancing test. The vendor also needs a lawful basis for any processing it does as controller (e.g. training), which is where consumer tiers fail: the Italian Garante fined OpenAI specifically for lacking a legal basis for training.

Special category data (Article 9). If the client data includes health, criminal convictions, biometric, or other special category data, processing is prohibited unless a specific exemption applies (e.g. legal professional privilege under Article 9(2)(e) in some member states). A local PII gate should fail-closed (block the request) when it detects Article 9 data.

Data minimisation (Article 5(1)(c)). The firm must process only the personal data adequate, relevant, and limited to what is necessary. Pasting an entire client file when the question is about one clause fails this principle.

International transfers (Articles 44-49, Chapter V). Transferring personal data outside the EEA (European Economic Area, the EU plus Iceland, Liechtenstein, and Norway) requires a legal basis: an adequacy decision (the Commission has not issued one for China or Singapore), Standard Contractual Clauses (SCCs), Binding Corporate Rules, or explicit consent. The EU-US Data Privacy Framework (DPF, 2023) provides a basis for transfers to certified US companies, but only when the US company is certified.

Breach notification (Articles 33-34). A personal data breach must be notified to the supervisory authority within 72 hours and to data subjects without undue delay if high risk. The Garante alleged OpenAI failed this duty for the 20 March 2023 ChatGPT incident. The later Rome judgment vacated the Italian sanction on competence; it did not bless a skipped notification.

Security of processing (Article 32). The firm and the processor must implement appropriate technical and organisational measures, including pseudonymisation. A local PII gate is an Article 32 measure.

DPIA (Article 35). A Data Protection Impact Assessment is required before high-risk processing, which includes large-scale processing of personal data and systematic evaluation of personal aspects. AI deployment on client data at scale triggers this.

Fines (Article 83). Up to EUR 20 million or 4 per cent of total worldwide annual turnover for the most serious violations (breaches of the principles in Articles 5, 6, 7, 9). Up to EUR 10 million or 2 per cent for less serious violations (Articles 28, 32, 33-34).

2.2 EU-level: The AI Act (Regulation (EU) 2024/1689)

The EU AI Act (the world's first comprehensive AI law) applies directly in Malta. It uses a risk-based framework:

- **Prohibited practices** (Article 5): social scoring, real-time biometric identification in public spaces (with narrow exceptions), manipulative AI. Not relevant to a law firm's internal use.
- **High-risk AI systems** (Annex III): AI used in the administration of justice, law enforcement, and biometric identification. A law firm using AI for document review is unlikely to be "high-risk" in the Annex III sense, but a court using AI for judicial decisions would be.
- **General-purpose AI models** (GPAI, Article 53): transparency obligations on model providers, applicable since August 2025. The law firm is a deployer, not a provider, so its obligations are lighter.
- **Transparency obligations** (Article 50): deployers of AI systems that interact with persons or generate content must inform users. A law firm using AI internally does not owe this to its lawyers, but if AI-generated content reaches a client deliverable, transparency may apply.

Malta implemented the AI Act through Legal Notice 226 of 2025 (under the MDIA Act, Chapter 591) and Legal Notice 227 of 2025 (under the Data Protection Act, Chapter 586). The MDIA (Malta Digital Innovation Authority) is the lead authority. The IDPC (Information and Data Protection Commissioner) has a specialised oversight role for AI systems implicating fundamental rights.

2.3 EU-level: EDPB guidance and enforcement

The European Data Protection Board (EDPB, the body of EU data protection authorities) has issued binding guidance on AI. Key positions:

- **No blanket AI exception to GDPR.** The EDPB has repeatedly confirmed that AI training and deployment must comply with all GDPR principles, including lawful basis, purpose limitation, and data minimisation.
- **No Article 22 exceptions.** Member states have not introduced exceptions to the prohibition on solely automated decision-making (Article 22) for AI. Data subjects retain the right to object.
- **Transfer risk assessments.** The EDPB requires transfer risk assessments (TRAs) for international data flows, considering the laws of the destination country (e.g. China's National Intelligence Law 2017, which requires companies to provide data access to government agencies).

2.4 Malta-specific: Professional secrecy and civil liability

This is the layer that makes a Malta law firm's situation stricter than a generic EU firm's.

Professional Secrecy Act (Chapter 377 of the Laws of Malta). Lawyers are bound by professional secrecy - the duty to keep client information confidential. This is broader than GDPR because it covers client confidential information regardless of whether it is "personal data." Disclosing client information to a third-party controller (e.g. consumer Claude) without client authority can breach this duty. The Ganado Advocates Chambers AI 2026 Malta guide confirms: "The use of AI, including generative AI tools, is governed by existing professional ethics rules, confidentiality obligations and data protection law. AI does not change the level of responsibility of lawyers to act ethically in accordance with the Code of Ethics that regulates the profession and their legal obligations resulting from, amongst other pieces of legislation, the Professional Secrecy Act (Chapter 377 of the Laws of Malta)."

Code of Organisation and Civil Procedure (Chapter 12). Governs court procedure and the legal profession in Malta. Reinforces the duty of confidentiality.

Article 1033 of the Civil Code (Chapter 16). Fault-based liability for damage caused by negligence, imprudence, or lack of attention. The Ganado guide notes: "The user cannot rely on ignorance of the effects of the use of the technology or the 'black box' phenomenon to avoid liability." So "I did not know the AI tool might leak client data" is not a defence.

No AI-specific guidance from the legal regulator (yet). The Committee for Advocates and Legal Procurators (within the Commission for the Administration of Justice) and the Chamber of Advocates have not issued AI-specific guidance for the legal profession. This means firms fall back on the existing frameworks above. Guidance is expected, but until it arrives, the duty is on the firm to interpret and comply.

Malta's AI-specific regulations (2025). Legal Notice 226 of 2025 designates the MDIA as Malta's lead authority for the AI Act. Legal Notice 227 of 2025 assigns the IDPC a specialised oversight role. No enforcement actions have been reported yet, but the framework is live.

The AI provider landscape: controller vs. processor by tier

The core distinction is not between Claude, Copilot, OpenAI, DeepSeek, and GLM. It is between the consumer tier and the commercial tier of each. The vendor's legal role flips based on which tier you use.

3.1 Anthropic (Claude)

Consumer tier (Claude Free, Pro, Max): - Anthropic acts as a **data controller**, not a processor. - **No DPA** is included. - Training is **opt-in** as of October 2025. If a user enables "Help improve Claude," retention extends to **five years**. - AMST Legal warns: "Using Consumer accounts for client-confidential work risks your data being stored for five years and used for training, which could violate professional secrecy or GDPR requirements." - Source: Anthropic Privacy Center, Janus Compliance, AMST Legal

Commercial tier (Claude for Work Team/Enterprise, Anthropic API, Claude Gov): - Anthropic acts as a **processor**. - "By default, we will not use your inputs or outputs from our commercial products to train our models." - The **DPA is built into the Commercial Terms** (current version applies from 1 January 2026). No separate signing step. - Default retention is **7 days** for API logs. **Zero Data Retention (ZDR)** is available for qualifying enterprise customers. - Source: Anthropic Privacy Center, Janus Compliance

3.2 OpenAI

Consumer tier (ChatGPT Free, Plus, Pro): - No DPA. Conversations **can be used for training** unless the account holder opts out. - Italy's Garante issued a EUR 15 million ChatGPT decision on 2 November 2024, including the finding that OpenAI had not identified a lawful basis for training. The Court of Rome annulled that decision on 18 March 2026 on one-stop-shop jurisdiction (Irish DPC as lead authority). The court did not decide the GDPR merits. - Source: Mondaq

Commercial tier (OpenAI API, ChatGPT Enterprise/Business): - OpenAI acts as a **processor** under its Data Processing Addendum. - DPA entered with **OpenAI Ireland Limited** for EEA and Swiss customers (Irish DPC is the lead authority). - API inputs and outputs are **not used for training** by default. - Retention: up to **30 days** for abuse monitoring. **Zero Data Retention (ZDR)** is approval-gated (request via Compliance area or api-compliance@openai.com). - **EU data residency** available for eligible customers: data at rest in Europe, in-region GPU inference in US or Europe. - Source: OpenAI Business Data, Janus Compliance, OpenAI Data Controls, Iubenda

3.3 Microsoft Copilot

Consumer tier (Copilot free, Copilot consumer): - Microsoft acts as controller. Not suitable for client-confidential work.

Enterprise tier (Microsoft 365 Copilot, Copilot Chat with enterprise tenant): - Microsoft acts as a **processor** under the Microsoft Products and Services DPA. - "Microsoft does not use enterprise tenant data to train its foundation models. This is a contractual commitment, not a best effort." - The DPA is **inherited from the existing M365 E3/E5 license** - no new contract required. - **EU Data Boundary (EUDB)** support: EU traffic stays within the EU for LLM processing. - Copilot respects existing M365 permissions (a user only sees content they could already access via SharePoint, Exchange, OneDrive). - Audit logs flow to the M365 unified audit log. - Source: Microsoft Learn - Enterprise Data Protection, Microsoft Learn - Privacy, AI Vortex Legal Guide

3.4 DeepSeek

Assessment: disqualifying for client-confidential work involving personal data.

- DeepSeek (Hangzhou DeepSeek Artificial Intelligence Co., Ltd. and Beijing DeepSeek Artificial Intelligence Co., Ltd.) is China-based.
- Data stored on servers in **China**, subject to Chinese intelligence laws (companies must provide data access to government agencies on request).
- **No adequacy decision** for China. DeepSeek's privacy policy does not mention SCCs or Binding Corporate Rules.
- **No EU representative** under Article 27 GDPR until late May 2025 - almost five months after the Italian ban.
- Italy's Garante imposed an **emergency ban** on January 30, 2025 (the first emergency ban on an AI chatbot under GDPR), citing violations of Articles 6, 12, 13, 14, 27, 31, and 32.
- South Korea's PIPC found DeepSeek had transferred user device information and **prompt content** to companies in China and the US (including Beijing Volcano Engine Technology) without consent. Transfers blocked from April 10, 2025.
- Source: Garante Provision, 30 January 2025, HIMSS, MIAI Regulation, Exterro

3.5 GLM / Zhipu AI (Z.ai)

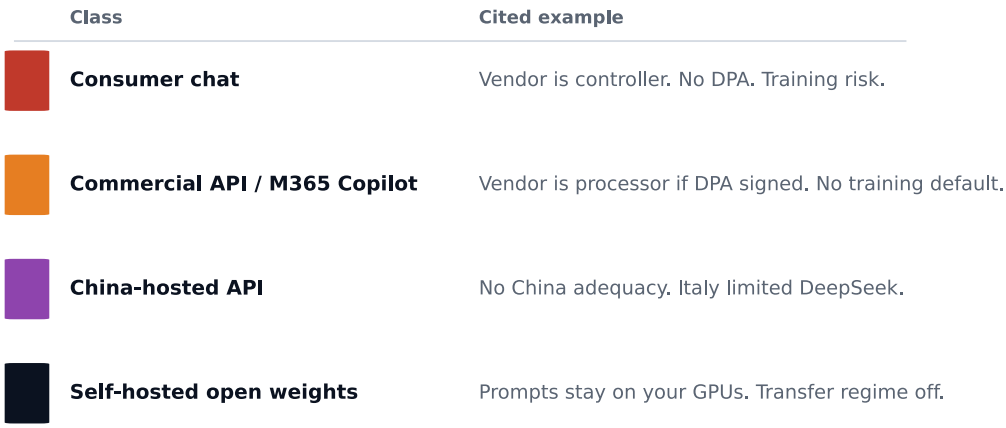
Assessment: the hosted API is disqualifying for personal data; self-hosting open weights is the only clean path.

- Zhipu AI is Beijing-headquartered, spun out of Tsinghua University in 2019.
- International API is contracted through a **Singapore-registered entity** (Jingsheng Hengxing Technology Pte. Ltd.).
- API terms state: real-time, non-persistent processing; not used for training under the DPA.
- BUT: **no adequacy decision** between the EU/UK and China or Singapore. The group remains under Chinese jurisdiction (National Intelligence Law 2017, Data Security Law).
- No EU data centers, no EU data residency, no EU SaaS offering documented.
- Most organisations cannot pass the transfer risk assessment required for personal data.
- **Self-hosting the MIT-licensed open-weight GLM models** on controlled infrastructure is the only method to avoid international transfer obligations entirely.
- As of June 2026, GLM-5.2 weights published under MIT licence.
- Source: Layer3Labs, Kifox, Security Scientist, Mondaq China

3.6 Summary matrix

Provider	Consumer tier	Commercial tier	DPA?	Training on data?	EU residency?	Adequate for client data?
Anthropic (Claude)	Controller, no DPA, opt-in training, 5yr retention	Processor, DPA in Commercial Terms, 7-day/ ZDR	Yes (commercial)	No (commercial)	Not stated	Commercial: Yes / Consumer: No
OpenAI	Controller, no DPA, training unless opted out	Processor, DPA with OpenAI Ireland, 30-day/ZDR	Yes (commercial)	No (commercial)	Yes (eligible)	Commercial: Yes / Consumer: No

Which path the prompt actually takes



Anthropic Privacy Center (controller vs processor); OpenAI Business Data; Microsoft DPA; Garante DeepSeek 30 Jan 2025; Reinvently GLM-5.2. AIMonger redraw.

Plain read: the legal question is the data path, not the brand name on the chat box. Consumer products put the vendor in a controller role. Commercial APIs and Microsoft 365 Copilot put the vendor in a processor role when a Data Processing Agreement is in force. China-hosted APIs fail Chapter V unless a transfer tool actually works. Self-hosted open weights keep prompts on infrastructure you control.

Source: Anthropic Privacy Center (controller versus processor, 16 March 2026); OpenAI Business Data; Microsoft Products and Services DPA; Garante DeepSeek provision 30 January 2025 [10098477]; Security Scientist GLM risk assessment; Layer3Labs Z.AI guide. AIMonger redraw. | Microsoft Copilot | Controller | Processor, M365 DPA inherited | Yes (enterprise) | No (enterprise) | Yes (EUDB) | Enterprise: Yes / Consumer: No | | DeepSeek | N/A (China-based) | China-based, no adequacy, no Art 27 rep | No | Unclear | No | No | | GLM (Zhipu) | China-based | Singapore entity, no adequacy | Yes (API) but no EU residency | No (API stated) | No | API: No / Self-hosted: Yes |

The evidence: court cases and data breaches

The user's instinct that "AI labs are not known to be careful with private data" is not speculation. It is documented in enforcement actions and court records.

4.1 OpenAI ChatGPT Redis data breach (March 2023)

On March 20, 2023, OpenAI took ChatGPT offline. A bug in the open-source library

```
redis-py
```

(the Python Redis client) caused a caching issue where some users could see other users' chat history titles. On deeper investigation, OpenAI discovered that 1.2 per cent of ChatGPT Plus subscribers active during a specific nine-hour window had payment-related information exposed: first and last names, email addresses, payment addresses, credit card type, last four digits of card numbers, and expiration dates. Full card numbers were never exposed. 440 Italian users were affected.

The root cause was a cancelled Redis request corrupting a shared connection, returning bad data for a different request. OpenAI added redundant checks and rate limiting. The bug was in a transitive dependency (redis-py), not OpenAI's own code.

Source: OpenAI Blog, 24 March 2023, The Verge, Safeguard Analysis

4.2 Italian Garante ChatGPT decision (2 November 2024) and Rome annulment (18 March 2026)

On 2 November 2024 the Garante adopted provvedimento n. 755 and calculated a total administrative sanction of EUR 15 million, plus a six-month Italian media campaign. The published breakdown included EUR 9 million for Articles 5, 6, 12, 13, 24 and 25 (lawful basis, transparency, age verification), EUR 320,000 for Article 33 (failure to notify the 20 March 2023 incident to the Garante), and EUR 5.68 million for Article 83(5)(e) (failure to run the earlier ordered information campaign). The Garante recorded 440 Italian users as affected by the March 2023 Redis incident.

On 18 March 2026 the Tribunale Ordinario di Roma (case R.G. 4785/2025) **annulled** that decision in full. Judge Damiana Colla held that after 15 February 2024 the Irish Data Protection Commission was OpenAI's lead supervisory authority under the GDPR one-stop-shop, so the Garante lacked competence to issue the November 2024 final decision. The court did **not** decide whether the alleged GDPR violations were proved. Wilson Sonsini, who represented OpenAI on the appeal, summarised the result as setting aside both the fine and the campaign order.

For a Malta firm this still matters. A regulator can open a file, allege missing lawful basis and missing breach notification, and force litigation. A later jurisdictional win for the vendor does not restore a client's professional-secrecy position if the firm's lawyers already pasted files into a consumer product.

Source: Garante provvedimento 2 November 2024, Rome judgment PDF, PPC.land, Wilson Sonsini, European Law Blog

4.3 Italian Garante emergency ban on DeepSeek (January 2025)

On January 30, 2025, the Italian Garante imposed an immediate ban on DeepSeek's data processing, citing violations of GDPR Articles 6 (lawful basis), 12 (transparency), 13-14 (information obligations), 27 (EU representative), 31 (cooperation with the authority), and 32 (security). This was the **first emergency ban on an AI chatbot under GDPR**, setting a precedent for preemptive action before a confirmed breach. DeepSeek's reply to the Garante's information request was treated as insufficient. The companies said they had not entered the Italian market and that the GDPR did not apply. The Garante disagreed and limited processing anyway.

Source: Garante Provision, 30 January 2025, HIMSS

4.4 South Korea PIPC investigation of DeepSeek (2025)

South Korea's Personal Information Protection Commission (PIPC) found that DeepSeek had transferred personal information (device/network/app information and the content of users' prompts) to multiple companies in China and the United States "for purposes such as service improvement and security" without obtaining user consent or disclosing the cross-border transfers. The transferred data included prompt content routed to Beijing Volcano Engine Technology. DeepSeek blocked new transfers from April 10, 2025, suspended new app downloads in Korea, and agreed to implement opt-out mechanisms.

Source: MIAI Regulation, RFA

4.5 Bartz v. Anthropic: the books conundrum (2024-2025)

In August 2024, authors Andrea Bartz, Charles Graeber, and Kirk Wallace Johnson filed a class-action lawsuit alleging Anthropic infringed copyrights by using **pirated books** (downloaded from shadow libraries including LibGen, Books3, and Pirate Library Mirror) to train its Claude LLMs. Anthropic downloaded over seven million pirated books between 2021 and 2022.

In June 2025, Judge William Alsup of the U.S. District Court for the Northern District of California ruled: - Training on **lawfully acquired** books is fair use ("quintessentially transformative"). - Downloading **pirated** copies to build a permanent "central library" is **not** fair use (infringement).

In August 2025 the parties reached a USD 1.5 billion settlement, preliminarily approved on 25 September 2025. On 20 July 2026 U.S. District Judge Araceli Martínez-Olguín granted **final approval**, describing it as the largest known U.S. copyright class-action settlement. Counsel fees were set at about USD 101.6 million (about 6.8 per cent of the fund). Payments to class members were expected to start after the appeal window (generally 30 days from the order). Reuters reported the same final-approval date.

The relevance to a law firm's risk assessment: this case proves that a leading AI lab (Anthropic, the developer of Claude) engaged in mass data acquisition from illegal sources without legal basis. The firm that assumes "the vendor is careful with data" is assuming against documented evidence. Anthropic's liability stemmed from the unlawful acquisition of source materials, not from the act of training per se, but the incident shows that data provenance and handling discipline at AI labs is not a given.

Source: Bartz v. Anthropic PBC, 787 F. Supp. 3d 1007 (N.D. Cal. 2025), Court Order on Fair Use, Buchanan Ingersoll & Rooney, Authors Guild, Publishers Weekly

4.6 The pattern

These incidents are not isolated. They form a pattern:

1. **Data breaches happen.** OpenAI's Redis bug exposed payment data. This was a software dependency issue, not a sophisticated attack, which means it can happen again to any provider.
2. **Training and acquisition disputes are public.** The Garante alleged ChatGPT training lacked a lawful basis. That allegation was not tried on the merits in Rome. Anthropic's pirated-book library was litigated and then settled for USD 1.5 billion with final approval on 20 July 2026. Consumer Claude still offers opt-in training with long retention if the lawyer turns improvement on.
3. **Regulators act preemptively.** The Italian Garante banned DeepSeek before a confirmed breach, setting a precedent that opaque data handling alone is grounds for action.
4. **Money is material even when a fine later falls.** The Garante still issued a EUR 15 million ChatGPT decision. Rome later set it aside on competence. Anthropic's USD 1.5 billion copyright settlement has final court approval. GDPR Article 83 still lists EUR 20 million or 4 per cent of worldwide turnover as the upper band for the most serious controller failures.
5. **China-based providers carry structural risk.** Chinese intelligence laws require companies to provide data access to government agencies. No adequacy decision exists for China. This is not a policy stance; it is a legal fact under GDPR Chapter V.

4.7 Another leadership team might argue the opposite

A partner could say commercial APIs already solve this, so a local gate is theatre. That argument is half true. A processor-tier contract plus Zero Data Retention is the minimum GDPR Article 28 story. It does not stop a lawyer pasting a name the model still needs for reasoning, and it does not convert a China-hosted API into an adequacy decision. The competing view is worth keeping as a check against over-engineering, not as a licence to skip the DPA.

A second counter-position is that Rome's OpenAI judgment proves EU enforcement is toothless. The judgment proves something narrower: once a main establishment sits in Ireland, other DPAs can lose

competence to issue a final cross-border fine. It does not erase Article 83, Malta's IDPC, Chapter 377 professional secrecy, or a client's complaint letter.

The runaway decisions: how small choices cascade into an audit

The path from "one lawyer tries ChatGPT" to "the IDPC opens an investigation" is not a single leap. It is a cascade of small, individually reasonable decisions that compound into a compliance disaster. Here is the cascade map.

5.1 The shadow AI cascade

Step	Decision	Consequence	Audit trigger
1	A lawyer pastes a client contract into consumer ChatGPT to summarise a clause	Client data is now in OpenAI's systems as a controller. No DPA, training opt-in possible.	None visible to the firm
2	The lawyer finds it useful and tells colleagues	Multiple lawyers adopt consumer ChatGPT for client work	Still invisible to IT and DPO
3	A lawyer pastes a contract with special-category data (e.g. a medical report in a personal injury case)	Article 9 data is now in a third-party controller's system without lawful basis	None visible
4	A client asks "how did you produce this summary so fast?"	The lawyer mentions AI. The client asks whether their data was used.	Client complaint to the IDPC
5	The IDPC requests the firm's records of processing (Article 30) and DPIA	No records exist because the use was informal	Article 30 and 35 violation
6	The IDPC discovers consumer tier use, no DPA, no lawful basis	Breach of Articles 5, 6, 28, 35	Fine: up to EUR 20M or 4% global turnover
7	The client sues for professional secrecy breach	Chapter 377 violation, Article 1033 civil liability	Malpractice claim, insurance impact
8	The matter becomes public	Reputational damage, other clients audit their files	Full regulatory investigation

5.2 The DeepSeek cost cascade

Step	Decision	Consequence	Audit trigger
1	IT adopts DeepSeek API because it is cheaper than Claude/ OpenAI	Client data flows to China. No adequacy, no SCCs, no Art 27 rep.	None visible
2	A lawyer uses DeepSeek for a matter involving an EU data subject	International transfer without legal basis (Chapter V violation)	None visible
3	The Garante (or IDPC) issues a ban on DeepSeek (as Italy did in Jan 2025)	All processing must stop immediately. Active matters are disrupted.	Enforcement order
4	The firm cannot recover data already sent	Data is in China, subject to Chinese intelligence laws. No deletion rights enforceable.	Article 17 (right to erasure) cannot be satisfied
5	Clients are notified that their data was sent to a banned provider	Breach notification obligation (Articles 33-34)	Class action exposure

5.3 The “no DPIA” cascade

Step	Decision	Consequence	Audit trigger
1	The firm adopts Claude API with DPA but skips the DPIA (“it is just an API”)	Article 35 violation. The firm cannot evidence that it assessed the risks.	None visible
2	An output hallucination reaches a client deliverable without review	The client receives incorrect legal advice. Malpractice exposure.	Client complaint
3	The IDPC reviews the firm’s AI governance	No DPIA = no evidence of risk assessment = higher fine	Article 35 violation fine
4	Article 1033 civil liability attaches	“The user cannot rely on ignorance of the effects of the technology“	Personal liability for the lawyer

5.4 The cascade principle

The pattern across all three cascades is the same: the first decision looks trivial (paste a contract, use a cheaper API, skip a document), and each subsequent step makes the situation worse and harder to reverse. By the time an audit triggers, the firm is dealing with multiple simultaneous violations (no DPA, no DPIA, no Article 30 records, Article 9 breach, international transfer violation, professional secrecy breach) and the defence of “we did not know” is explicitly unavailable under Article 1033.

The lesson: the decision to gate AI use - to control which tools, which tiers, and with what safeguards - must happen before the first lawyer opens a chat box, not after the IDPC sends a letter.

Mitigation: the local PII gate

This is the architectural defence that no contract clause can match. The idea is simple: run a local model that intercepts every prompt before it reaches a cloud API, detects personal data, and replaces it with semantic placeholders. The cloud model never sees the real PII (Personally Identifiable Information). On the return path, the gate restores the original values.

6.1 The architecture

```
[Lawyer's prompt]
      |
      v
[Local PII Gate]
  1. Regex layer: emails, phone numbers, IBAN, IDs, card numbers (deterministic, zero
latency)
  2. NER layer: person names, organisations, locations (local model, CPU)
  3. LLM layer: context-dependent entities, trade secrets, Art 9 data (local LLM via
Ollama)
      |
      v
[Pseudonymised prompt] -- sent to cloud API (Claude / OpenAI / Copilot)
      |
      v
[Cloud LLM response]
      |
      v
[Local PII Gate: reverse substitution using encrypted mapping table]
      |
      v
[Restored response with real values] -- returned to lawyer
```

Three controls before a client file leaves the firm

GDPR Arts 28, 32, 35. Local-gate projects: Hey Jude, Septum, pii-proxy. AIMonger framework map (no scores).

On Monday: a managing partner can ask whether the firm has all three layers, or only a licence invoice. A Data Processing Agreement without a local gate still sends identifiers if lawyers paste them. A gate without a processor-tier contract still leaves residual text with a vendor whose terms you have not locked.

Source: GDPR Articles 28, 32 and 35; Hey Jude, Septum and pii-proxy project documentation. AIMonger framework map (no invented scores).

6.2 Why this changes the legal analysis

The gate is not just a technical convenience. It changes the GDPR analysis:

1. **Pseudonymisation (Article 32(1)(a)).** Replacing PII with placeholders is pseudonymisation, which is an explicit GDPR security measure. The firm can evidence Article 32 compliance by showing the gate is in place.
2. **No personal data leaves the EEA.** If the pseudonymised prompt contains no personal data, the international transfer regime (Chapter V, Articles 44-49) does not trigger. This is the single most powerful legal effect: it removes the transfer risk for DeepSeek and GLM, and reduces the transfer risk for US providers.
3. **Fail-closed on Article 9.** If the gate detects special-category data (health, religion, biometric), it blocks the request. This prevents the most serious GDPR violation before it occurs.
4. **Article 30 records.** The gate's audit log (JSONL format: what was sent, to which LLM, when, by whom) is the data source for the firm's records of processing activities.
5. **Re-identification resistance.** A good gate does not just strip the literal name; it tests whether the sanitised text can still re-identify the subject through context. Hey Jude runs a "re-identification critic" - a blind attack on the sanitised output - and broadens any placeholder a blind attacker could pin down.

6.3 Open-source tools

Tool	Focus	Key features	Best for
Hey Jude	Legal workflows	Local LLM swap, Presidio safety net, re-identification critic, jurisdiction-aware routing, IRREDUCIBLE_POLICY (block/ask/route to sovereign model)	Law firms with mixed sensitivity matters
Septum	Document-heavy workflows	Three-layer detection (Presidio + NER + Ollama), approval gate, audit trails, 7 modules across 3 security zones, air-gapped PII handling	Firms processing large document sets
pii-proxy	GDPR compliance	Regex + local LLM, blocks Art 9 data, AES-256-GCM mapping, JSONL audit log for Art 30, fail-closed	EU firms needing Article 30 evidence
localscrub	Clinical/free-form text	Two-stage cascade (rules/NER + local LLM), all 18 HIPAA Safe Harbor categories, air-gapped	Firms handling medical/criminal law matters

6.4 The gate is necessary, not sufficient

The local PII gate is a powerful defence, but it does not replace the other obligations. The firm still needs: - A processor-tier product (the gate pseudonymises, but the cloud model still processes the pseudonymised text - the DPA still matters for the residual data). - A DPA in place (for the non-PII data that does flow, and for the mapping table if stored). - A DPIA (the gate is a risk mitigation measure, not a substitute for the assessment). - Human oversight (the gate handles data protection; it does not handle hallucination, accuracy, or privilege review).

Cost analysis

The cost of compliant AI is not just the API bill. It is the sum of the product tier, the DPA overhead, the DPIA time, the gate infrastructure, and the operator cost.

7.1 Tier comparison

Path	Monthly cost (per user or per node)	Compliance posture	Notes
Consumer Claude/ ChatGPT	~$20/user	Non-compliant for client data	Controller role, no DPA, training risk
Claude API (Sonnet)	~$3/$15 per 1M tokens in/out	Compliant with DPA + DPIA	Processor role, 7-day retention, ZDR available
OpenAI API	~$5/$30 per 1M tokens (premium)	Compliant with DPA + DPIA	Processor role, 30-day retention, EU residency available
M365 Copilot (enterprise)	~$30/user/month	Compliant (DPA inherited)	Processor role, EUDB, audit logs in M365
DeepSeek API	Cheapest ($0.14/$0.28 per 1M)	Non-compliant for EU personal data	China-based, no adequacy, banned in Italy
GLM self-hosted (open weights)	GPU cost only	Compliant (no transfer)	MIT licence, self-host on EU GPUs
Local PII gate (Hey Jude/ Septum/pii-proxy)	GPU for local model + open source	Adds Article 32 defence	Run on existing Mac Studio or rented GPU

7.2 The gate cost

A local PII gate runs a small local LLM (typically 7B-13B parameters via Ollama or LM Studio) for PII detection. This can run on: - A Mac Studio M2 Ultra (already common in law firms for document review): no additional hardware cost. - A rented A100/H100 GPU if higher throughput is needed: approximately $2-4/GPU-hour on spot, $6-8 on-demand. - The gate does not need a frontier model; a 7B model is sufficient for NER and context-dependent PII detection.

7.3 The DPIA cost

A DPIA (Article 35) for AI deployment is a structured document that a DPO or external advisor produces. It typically takes 2-5 days of professional time for a first deployment, less for subsequent similar deployments. This is a one-time cost per use case, not a recurring per-query cost.

7.4 The full sovereign path

For firms that cannot accept any data leaving their boundary (e.g. top-tier litigation, M&A with inside information), the alternative is self-hosting an open-weight model (GLM, Llama, or DeepSeek's open weights) on dedicated EU-region GPUs. The cost is dominated by always-on GPU capacity: approximately $39,700/month for a full B200 tray (8 GPUs at $6.89/GPU-hour, 720 hours/month), per the cost analysis in AIMonger's Cost of Privacy whitepaper. This is the privacy premium for full sovereignty - expensive but justified only for the most sensitive matters.

7.5 The practical recommendation

For most Malta law firms, the cost-effective compliant path is: 1. **Claude API or OpenAI API** for the cloud model (processor role, DPA, no training). 2. **A local PII gate** (Hey Jude or pii-proxy) running on an existing Mac Studio or modest GPU. 3. **M365 Copilot** for lawyers who need integrated document AI (DPA inherited, EUDB, audit logs). 4. **A DPIA** per material use case. 5. **Reserve self-hosted open weights** for matters where no data may leave the boundary.

This path costs approximately $30-50 per lawyer per month in API usage, plus a one-time DPIA investment, plus the existing hardware for the gate. It is materially cheaper than the cost of a single GDPR fine or a single malpractice claim.

Decision framework: the compliance checklist

#	Requirement	Source	How to verify
1	Use a processor-tier product, not consumer	GDPR Art. 28	Check account type: API/Enterprise/Team, not Free/Pro/Max
2	DPA signed and on file	GDPR Art. 28(3)	Anthropic: Commercial Terms. OpenAI: platform.openai.com > Compliance. Microsoft: inherited from M365 license
3	No training on your data	GDPR Art. 5(1)(b)	Confirm in vendor terms: commercial tier says “not used for training”
4	EU data residency (if using US provider)	GDPR Chapter V	OpenAI: request EU residency. Microsoft: confirm EUDB. Anthropic: request ZDR
5	DPIA completed for each material use case	GDPR Art. 35	Document on file; review annually
6	Article 30 records of processing	GDPR Art. 30	PII gate audit log + manual register
7	Local PII gate installed	GDPR Art. 32(1)(a)	Hey Jude / Septum / pii-proxy running locally
8	Fail-closed on Article 9 data	GDPR Art. 9	Configure gate to block, not just pseudonymise, special-category data
9	Re-identification critic enabled	GDPR Art. 32	Hey Jude REID_CRITIC_ENABLED = true
10	Human review of every AI output	AI Act + professional ethics	Policy: no AI output to client deliverable without lawyer review
11	Audit trail of who used what, when	AI Act + Art. 1033 Civil Code	Gate log + M365 audit log
12	No consumer AI on client data (shadow AI policy)	Professional Secrecy Act Ch. 377	Written policy + periodic device/app audit
13	No China-based providers for personal data	GDPR Chapter V	Block DeepSeek and GLM hosted API at network level; allow self-hosted GLM open weights
14	Client engagement terms address AI use	GDPR Art. 6(1)(a)	Update engagement letters: disclose AI use, obtain consent where needed
15	Breach response plan covers AI incidents	GDPR Art. 33-34	Test the plan: what happens if the gate fails or the vendor has a breach?

What I have not verified

- I have not read the full text of Anthropic's DPA, OpenAI's DPA, or Microsoft's DPA. The analysis relies on the vendors' own published summaries and secondary analyses (Janus Compliance, AMST Legal, Iubenda, Compound Law).
- I have not confirmed whether the Committee for Advocates and Legal Procurators has issued AI-specific guidance since the Ganado guide was published (June 2026). The Ganado guide states none had been issued as of their writing.
- I have not verified a specific firm's M365 tenant configuration or whether the EU Data Boundary is enabled - that is a tenant-specific check.
- I have not priced the exact token consumption for a typical law firm matter. The API costs above are list prices per million tokens; actual usage depends on document length and query frequency.
- This dossier is information and analysis, not legal advice. For a binding sign-off, a Malta-qualified data protection advisor (the IDPC publishes a register of DPOs and advisors) must review the firm's specific engagement letters, IT configuration, and matter types before confirming compliance.

Conclusion

The instinct that "running Claude or Copilot on private information creates a legal constraint" is correct. The constraint is not the tool; it is the data flow. Consumer tiers make the vendor a controller of client data with no DPA and possible training use - a position that breaches GDPR and Malta's Professional Secrecy Act. Commercial tiers reframe the vendor as a processor with a DPA and no training, which handles most of the GDPR mechanics contractually. A local PII gate adds a structural defence by pseudonymising prompts before they leave the firm's boundary, satisfying Article 32 and removing the international transfer risk.

The riskiest move is the one most lawyers default to: pasting into the consumer chat box they already have open. The compliant path (processor-tier product, DPA, DPIA, local gate, human review) is not free, but it is cheaper than a surviving GDPR fine, a Chapter 377 complaint, or a single professional secrecy claim.

Malta's position as an EU member state with a specific professional secrecy regime (Chapter 377) and a fault-based civil liability standard (Article 1033, with no "I didn't know" defence) makes the stakes higher, not lower, than for a generic EU firm. The framework is now live: Legal Notice 226 and 227 of 2025 implement the AI Act, the MDIA and IDPC are the regulators, and the Italian Garante's enforcement against OpenAI and DeepSeek shows that EU data protection authorities act, not just warn.

The technology to comply exists. The legal framework is clear. The court cases prove the risk. The remaining question is whether the firm's leadership treats this as a decision to make before the first lawyer opens a chat box, or a problem to explain after the IDPC sends a letter.

References

EU legislation

1. Official Journal of the European Union. "Regulation (EU) 2016/679 of the European Parliament and of the Council" (GDPR) - https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679

2. Official Journal of the European Union. “Regulation (EU) 2024/1689 of the European Parliament and of the Council” (AI Act) - https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Malta legislation

3. Legislation Malta. “Professional Secrecy Act” (Chapter 377) - https://legislation.mt/eli/cap/377/eng

4. Legislation Malta. “Code of Organization and Civil Procedure” (Chapter 12) - https://legislation.mt/eli/cap/12/eng

5. Legislation Malta. “Civil Code” (Chapter 16) - https://legislation.mt/eli/cap/16/eng

6. Legislation Malta. “Data Protection Act” (Chapter 586) - https://legislation.mt/eli/cap/586/eng

7. Legislation Malta. “Malta Digital Innovation Authority Act” (Chapter 591) - https://legislation.mt/eli/cap/591/eng

8. Legislation Malta. “Artificial Intelligence Regulations, 2025” - https://legislation.mt/eli/ln/2025/226

9. Legislation Malta. “Intelliġenza Artifiċjali” - https://legislation.mt/eli/ln/2025/227

Provider documentation

10. Anthropic Privacy Center: Controller vs. Processor - https://privacy.claude.com/en/articles/9267385-does-anthropic-act-as-a-data-processor-or-controller

11. Anthropic Privacy Center: Model Training - https://privacy.claude.com/en/articles/7996868-is-my-data-used-for-model-training

12. Anthropic. “How do I view and sign your Data Processing Addendum (DPA)?” - https://privacy.claude.com/en/articles/7996862-how-do-i-view-and-sign-your-data-processing-addendum-dpa

13. OpenAI Business Data Privacy - https://openai.com/business-data/

14. OpenAI Data Controls - https://developers.openai.com/api/docs/guides/your-data

15. OpenAI March 20 Outage Post - https://openai.com/blog/march-20-chatgpt-outage

16. Microsoft Copilot Enterprise Data Protection - https://learn.microsoft.com/en-us/microsoft-365/copilot/enterprise-data-protection

17. Microsoft Copilot Privacy and Protections - https://learn.microsoft.com/en-us/copilot/privacy-and-protections

18. Microsoft 365 Copilot Privacy - https://learn.microsoft.com/en-us/microsoft-365/copilot/microsoft-365-copilot-privacy

Court cases and enforcement actions

19. Bartz v. Anthropic PBC, 787 F. Supp. 3d 1007 (N.D. Cal. 2025) - https://www.copyright.gov/fair-use/summaries/Bartz-v-Anthropic-PBC-787-F-Supp-3d-1007-ND-Cal-2025.pdf

20. Bartz v. Anthropic Court Order on Fair Use - https://storage.courtlistener.com/recap/gov.uscourts.cand.434709/gov.uscourts.cand.434709.231.0_4.pdf

21. Italian Garante Provision v. DeepSeek, 30 January 2025 [10098477] - https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/10098477

22. Garante ChatGPT decision 2 November 2024 - https://www.civile.it/privacy/visual.php?num=99128

23. Tribunale di Roma, OpenAI v Garante (R.G. 4785/2025), 18 March 2026 - https://dei.web.uniroma1.it/sites/default/files/allegati/2026-05/Trib_Roma_OpenAI_Garante_2026.pdf

24. Reuters: US judge approves Anthropic USD 1.5 billion settlement, 20 July 2026 - https://www.reuters.com/world/us-judge-approves-anthropics-15-billion-settlement-copyright-lawsuit-2026-07-20/

Secondary analysis

25. Janus Compliance: Is the Claude API GDPR Compliant? (2026) - https://www.januscompliance.co.uk/blog/is-the-claude-api-gdpr-compliant-2026

26. Janus Compliance: ChatGPT / OpenAI DPA Explained (2026) - https://www.januscompliance.co.uk/blog/openai-dpa-data-processing-agreement-2026

27. AMST Legal: Anthropic's Claude AI Updates - https://amstlegal.com/anthropics-claude-ai-updated-terms-explained/

28. Iubenda: OpenAI GDPR Compliance in 2026 - https://www.iubenda.com/en/blog/openai-gdpr-compliance/

29. Compound Law: OpenAI API GDPR Analysis - https://compound.law/en-DE/tools/openai-api/

30. AI Vortex: Microsoft Copilot Attorney-Client Privilege Guide (2026) - https://www.aivortex.io/legal/guides/microsoft-copilot-attorney-client-privilege-confidential-data-2026/

31. Wilson Sonsini: OpenAI prevails in Italian GDPR case, 31 March 2026 - https://www.wsgr.com/en/insights/openai-prevails-in-landmark-italian-ai-and-gdpr-enforcement-case.html

32. PPC.land: Rome court annulment of EUR 15M OpenAI fine - https://ppc.land/italian-court-kills-openais-eur15m-fine-and-it-wasnt-even-close/

33. Mondaq: Artificial Intelligence in Malta - https://webiis10.mondaq.com/new-technology/1710984/artificial-intelligence-in-malta

34. Ganado Advocates, Chambers AI 2026 Malta - https://ganado.com/wp-content/uploads/2026/06/Chambers_Artificial-Intelligence-2026_016_malta.pdf

35. MIAI Regulation: DeepSeek One Year Later - https://ai-regulation.com/deepseek-one-year-later-regulatory-storm-global-surge/

36. HIMSS: DeepSeek Blocked in Italy - https://www.himss.org/news-center/deepseek-blocked-italy-due-privacy-risks-setting-significant-precedent/

37. Exterro: DeepSeek Data Privacy Alert - https://www.exterro.com/resources/data-privacy-alerts/data-privacy-alert-global-scrutiny-over-deepseeks-data-practices-intensifies

38. Security Scientist. "Zhipu GLM Risk Assessment Template" - https://www.securityscientist.net/blog/zhipu-glm-risk-assessment-template/

39. Layer3Labs: Z.AI Explained - https://www.layer3labs.io/guides/z-ai-explained

40. Kifox: Zhipu AI Z.AI GLM - https://kifox.ai/en/ki-tools/zhipu-ai-z-ai-glm-en/

41. Mondaq: Legal Considerations for EU Businesses Using China-Based AI Services - https://www.mondaq.com/china/data-protection/1790384/legal-considerations-for-eu-businesses-using-china-based-ai-services-part-ii-personal-data-protection

42. Buchanan Ingersoll & Rooney: Anthropic Copyright Settlement - https://www.bipc.com/anthropic%E2%80%99s-copyright-settlement-lessons-for-ai-developers-and-deployers

43. Authors Guild: Bartz v. Anthropic Settlement - https://authorsguild.org/advocacy/artificial-intelligence/what-authors-need-to-know-about-the-anthropic-settlement/

44. Publishers Weekly: Federal Judge Rules AI Training Is Fair Use - https://www.publishersweekly.com/pw/by-topic/digital/copyright/article/98089-federal-judge-rules-ai-training-is-fair-use-in-anthropic-copyright-case.html

45. IBA: Balancing Efficiency and Privacy: AI's Impact on Legal Confidentiality - https://www.ibanet.org/balancing-efficiency-and-privacy-AI-impact-on-legal-confidentiality-and-privilege

46. Kiteworks: AI Compliance for Legal - https://www.kiteworks.com/regulatory-compliance/ai-compliance-legal-departments-law-firms/

Local PII gate tools

47. GitHub. "sure-scale/hey-jude" - https://github.com/sure-scale/hey-jude

48. Septum - https://github.com/byerlikaya/Septum

49. pii-proxy (whitestag-ai) - https://github.com/whitestag-ai/pii-proxy

50. pii-proxy (daslabhq) - https://github.com/daslabhq/pii-proxy

51. localscrub - https://github.com/valbarov/localscrub

Published by AIMonger . https://aimonger.com . Malta . Europe