



amazee.io
part of Mirantis

Your AI Decisions Are Creating Risk: What Agency and Business Leaders Need to Own Now

Data Sovereignty as the Foundation for
Trustworthy AI on the Open Web

Asheville: July 11 2026



Matthew Saunders

- 20+ years in Drupal
- amazee.io AI Ambassador
- Open Source Maximalist
- matthews on Drupal.org



AI Honeymoon

- **People in love give everything**
 - Secrets, keys, blind trust
 - No second thoughts, consequences don't exist yet
- **Organizations are in that exact phase with AI**
 - Uploading contracts, source code, customer records, strategy docs
 - No hesitation, no questions, pure infatuation
- **Every honeymoon ends**
 - Where did my data go? Who has access? What do the terms actually say?
 - You gave away more than you should have
- **This talk is about what comes next**
 - What happens when the honeymoon is over
 - How to build a relationship with AI that you actually control

AI's Data Control Crisis

- **Samsung source code leak (2023)**

Engineers uploaded sensitive source code into an unauthorized AI tool, a textbook shadow AI incident

- **IBM 2025 Cost of a Data Breach Report**

20% of orgs suffered a shadow AI breach, costing on average \$670K more than traditional incidents. 77% of employees paste data into GenAI tools.

- **2025 Thales Data Threat Report**

69% of organizations cite AI-powered data leaks as their top security concern

- **Adversa AI 2025 Report**

AI security incidents doubled since 2024; 35% caused by simple prompt injection; some led to \$100K+ real losses



RTFM

Read The Flippin Manual



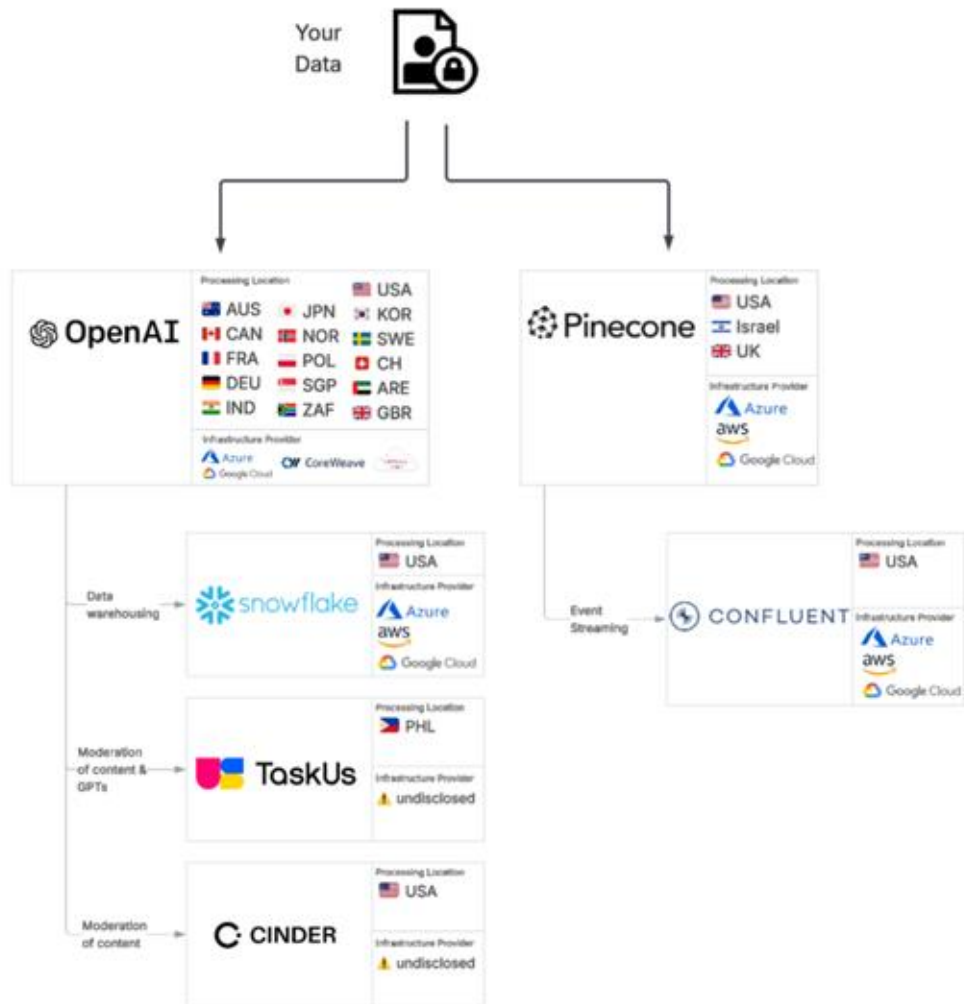
RTFPP

Read The Flippin Privacy Policy

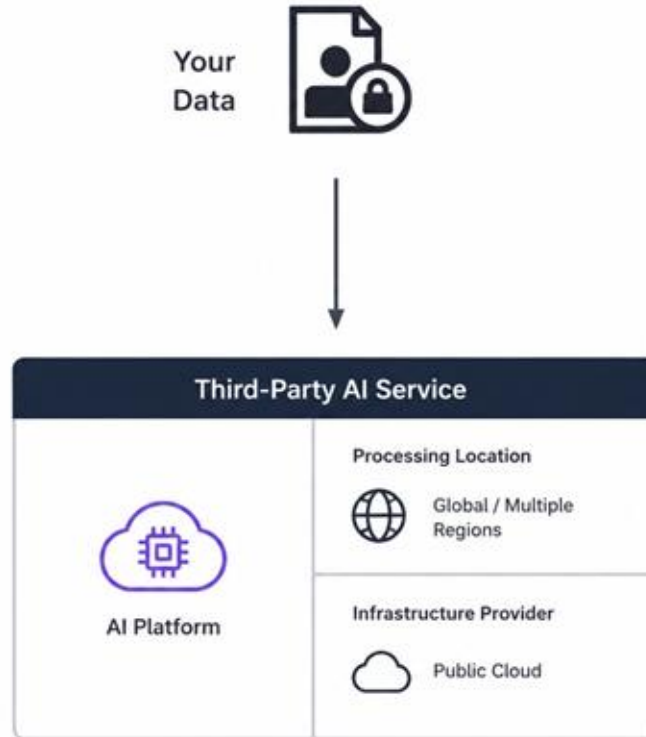
Privacy Policies

	OpenAI	Anthropic
Policy complexity	6+ documents, ~4,500 words each	4+ documents, ~4,000 words each
Data hosting countries defined?	Only for opt-in API customers	"Worldwide" — undefined
Subprocessor count	Large, locations often blank	10+ listed, mostly US-based
Default data residency	US	US
EU residency available?	Recently, API only	Recently announced, limited
Training on your data?	Opt-out for consumer plans	Opt-out since Aug 2025
CLOUD Act exposure?	Yes	Yes
Can change sub processors?	30-day notice, take it or leave	Similar mechanism

How it is



How it should be



Privacy Concerns with OpenAI & Anthropic

Data location: undefined

- Hosting listed as "Worldwide" or left blank
- No guarantees on which country processes your prompts

Long subprocessor chain

- 10+ third parties touch your data (billing, analytics, support, moderation)
- Mostly US-based → CLOUD Act applies

Opted into training by default

- Consumer plan data used for model training unless you actively disable it
- Most employees on free tiers never change the setting

Policy-based, not architecture-based privacy

- No per-customer isolation, no self-hosting, no audit trail
- Trust a legal document, not a technical architecture

The Three Pillars of AI Sovereignty

Data Sovereignty

- Your data, in your jurisdiction, under your control
- Regional data processing: CH, DE, UK, US, AU - data never leaves the chosen region
- Zero data retention: no prompt/output logging, metadata only

Model Sovereignty

- Multi-model flexibility: Claude, GPT, Mistral, open-source options; switch without changing application code
- No vendor lock-in: unified API abstracts individual model providers
- OpenAI-compatible API standard for interchangeability

Operational Sovereignty

- Not tied to a specific cloud provider
- Deploy on-premise, in-region cloud, or hybrid
- Audit-ready logging, role-based access controls out of the box
- Enterprise trust layer: ISO 27001, SOC 2, GDPR compliant

The Regulatory Landscape Demands This

- **EU AI Act now in force**
conformity assessments, risk management, human oversight required
- **GDPR enforcement escalating**
€530M fine on TikTok for transferring EEA user data to China
- **California expanding AI privacy bills and data broker disclosure requirements**
- **US CLOUD Act risk**
23% of Canadian orgs actively migrating away from US-headquartered cloud providers
- **1 in 3 organizations experienced a data sovereignty incident last year**

Open Source is Not a License

- **Transparency over secrecy**
- **Portability over lock-in**
- **Community over monopoly**
- **Accountability over blind trust**

Open Source is Not a License

- **Governments call this digital sovereignty**
- **CIOs call it governance**
- **Architects call it portability**

They're describing the same thing

Real-World Deployment Patterns

- **Private LLM proxies**

API normalization, token counting, budgeting, monitoring, NO logging

- **Region-locked LLM Providers**

Multiple Providers that offer LLMs to run region locked

- **Region-locked vector databases**

Content indexed through Search API, embeddings stored in dedicated per-customer databases

- **Guardrails**

Input-output filtering, prompt injection defenses

How Open Source + Drupal Make It Possible

Architecture walkthrough:

- LiteLLM proxy provides API normalization across LLM providers
- VectorDB (PostgreSQL + PgVector) for semantic search

Drupal AI Provider integration:

- The only Drupal AI provider bundling LLMs + vector database in one package
- Search API integration for granular content indexing control
- Single-click install, zero learning curve

Framework & Takeaways

Decision framework:

- If evaluating AI architecture
 - demand per-tenant isolation, regional deployment, zero logging
- If managing shadow AI risk
 - provide a sanctioned, private alternative (the Samsung lesson)
- If building AI features into Drupal
 - leverage Drupal AI Provider with sovereignty built in

Key takeaways:

- Isolation by default reduces risk.
- Open source provides governance and portability.
- The stack stays open, portable, and composable.

Thank You!

