CSA 2026:
The Cure for CEO FOMO

A Practical Guide to Enterprise AI Readiness

https://epsd.io

March 23, 2026

Presented by

Nick Selby, EPSD

Contributing Authors:

Pablo Breuer, EPSD

Huw Rogers, EPSD

Sarah Wells, EPSD

Chris Swan, At Sign

Table of Contents

A PDF of this presentation is available at the QR code below

00 The Basics

You’d be stunned how many firms don’t consider this.

01 What could possibly go wrong?

Strategy. Integration Control. Monitoring.

02 Modeling, not Narratives

Detecting, monitoring, blocking.

03 3PR and Blast Radius

Oops.

04 References

Q&A I’ll try.

Any Tech Can Make You Sad

This talk has lots of ways AI can make you sad. There are simply astoundingly great use-cases for AI.

- If you need someone to tell you about those, it may be hard ….
- It would be the height of hypocrisy for me to slag off AI: For years, I helped a billionaire sell magic internet money.
- People are often shocked, though, at the **basic things** I am about to tell you, so in the spirit of equal time….

I'm not an AI skeptic. This is not an anti-AI talk.

A PDF of this presentation is available at the QR code:

Executive Summary

This seems totally obvious but I cannot tell you how many companies skip one or more of these:

What exactly do you want AI to *do* for your business?

For your business, revenues, productivity? Specifically.

Modeling, not Narratives

Business use-cases, data security, licensing, must be clear to the CEO & board

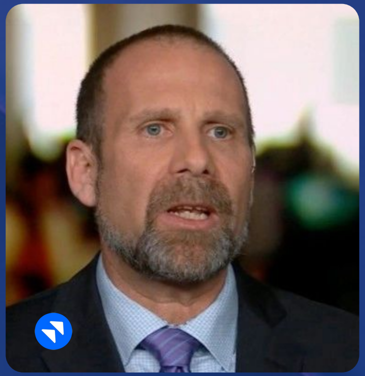
Blast Radius Reduction, Modeling, Alerting

If you can't see it, you can't fix it when it bursts into flames (and it will)

Nick Selby

Founder and Managing Partner at EPSD, Inc.

EPSD provides low-variance outcomes for high-growth executive teams.



About Me

30+ years of enterprise technology experience.

- **Private sector**: EPSD, Evertas Professional Services, Trail of Bits, Paxos, Bishop Fox, N4Struct/TRM Partners, 451 Research (now S&P Global Intelligence)
- **Public Sector**: Director of Cyber Intelligence & Investigations, NYPD; police detective investigator, DFW-area, 2010-2025
- **Co-author**: Cyber Survival Manual: From Identity Theft to The Digital Apocalypse; Blackhatonomics: An Inside Look at the Economics of Cybercrime, technical editor, Investigating Internet Crimes

00

Basics

Speak simply.

Say the quiet part aloud.

Repeat.

Disclaimer: I Shouldn’t Have To Say Any Of This…

Wear a hat when it’s cold.

Save cash for a rainy day.

Hydrate.

“Keep it simple. No, simpler.”

Like any tool, AI is not a strategy, it’s an untethered tactic. Without a clear and compelling strategy, AI is an expense with unclear ROI that, beyond known integration and third-party risks, carries unknown business, revenue, and liability perils.

In InfoSec strategy, use sentences like, “Help executives take better, and better-informed, risks to achieve better results using AI,” and “Defend against known bad; detect and stop previously unknown bad.”

Don’t make it more complex than it is.

“Love the boys. Explain the mission. Don’t be a detriment to the fighting.”

— LCDR Clint Bruce, SEAL Team 5, on SEAL team commander strategy.

“Upkeep, not just acquisition”

A very common situation is that the team buying or building a tool or application won’t be the team that owns and maintains it. Or, it’s given short shrift, and expected to be worked on “in your spare time.”

To ensure ongoing security: testing, patching, upgrading, and updating are essential. Infosec can’t do all that itself.

Good stewardship is essential to cyber security.

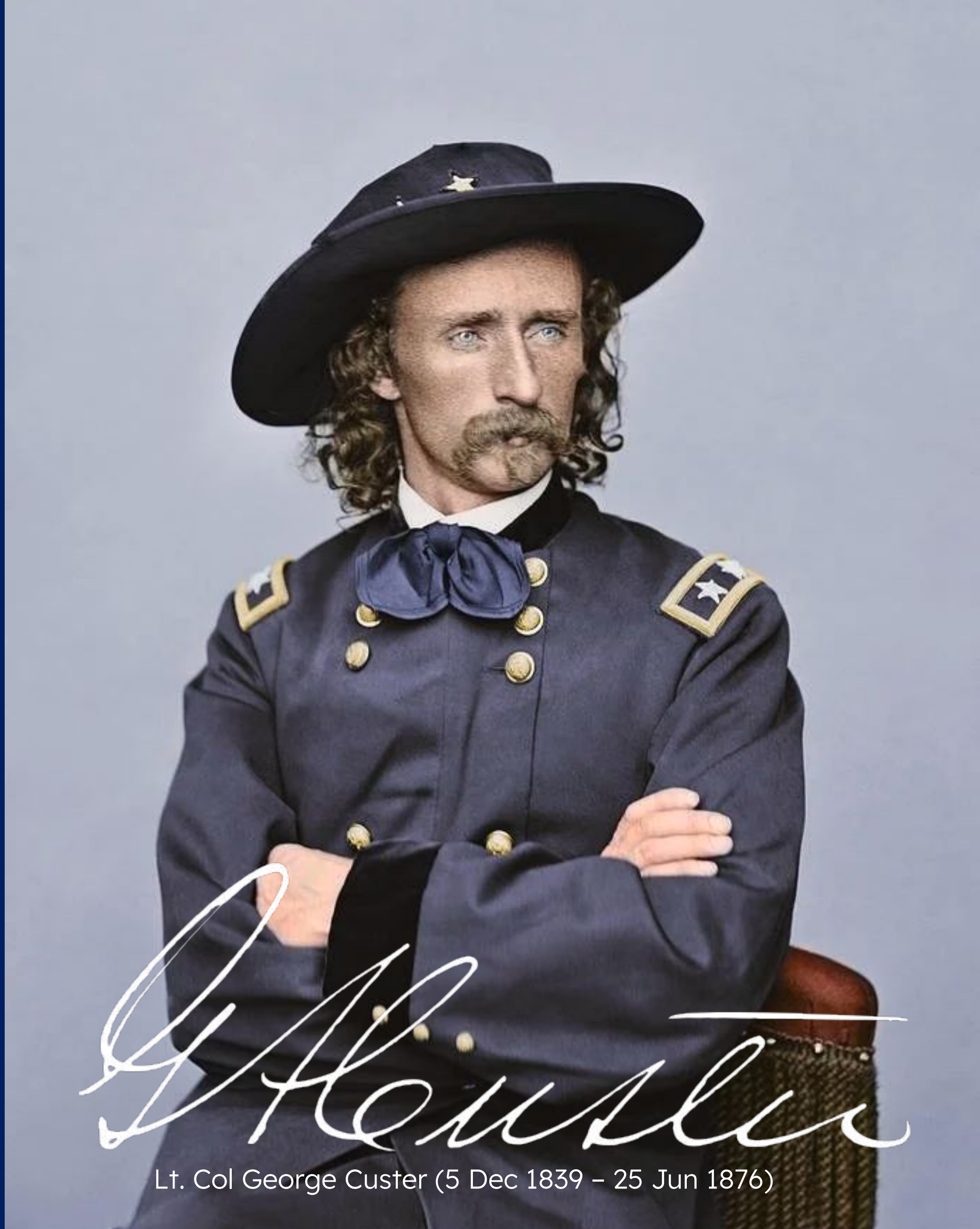
“Support and maintenance is part of running any software system; it’s the cost of doing business.”

— Sarah Wells, EPSD

Instead: “Let’s Goooooo!”

None of this has stopped CEOs from demanding that their organizations go “AI First”! There are legitimate reasons they feel like this:

- A commitment to advancing competitiveness
- An unbounded but genuine belief that AI tools increase “velocity”
- Fear Of Missing Out
- “Don’t give me excuses. AI!”
- Have we ever seen this before?



Lt. Col George Custer (5 Dec 1839 – 25 Jun 1876)

01

Define Your Terms

Chatbots

AI Tools

MCP Servers

AI Makers Have Cooked The Books

Khlaff and Myers West On Language Hacking

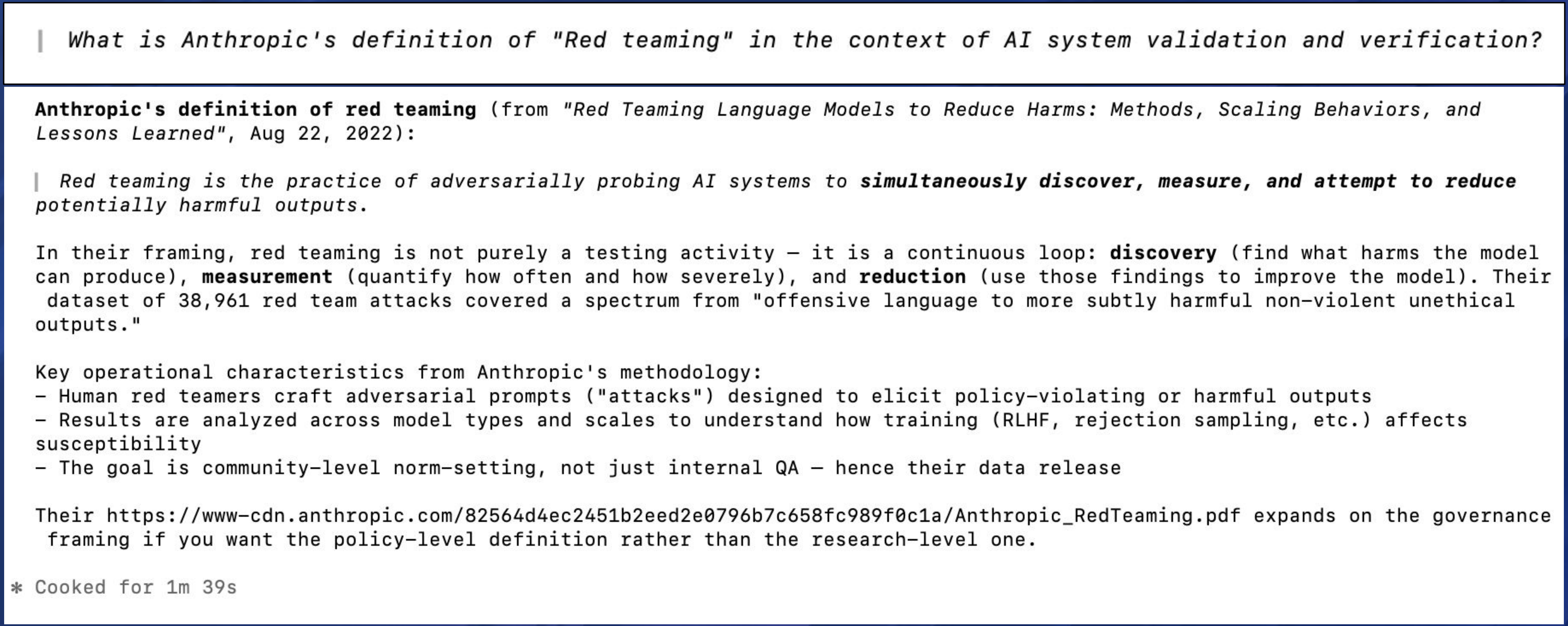
They state that "AI safety" discourse has misappropriated terminology of safety engineering, and verification and validation to make unverifiable claims about AI capabilities, redefining technical and regulatory terms in ways that undermine assurance methodologies, and compromise safety-critical and defense systems.

“AI Safety” → “[O]ften hinges on unfounded assumptions regarding the capabilities and “intent” of AI systems that may be leveraged to justify the deployment of risky systems.

Read the work → Safety Co-Option and Compromised National Security: The Self-Fulfilling Prophecy of Weakened AI Risk Thresholds

- Heidy Khlaaf, Sarah Myers West
- https://doi.org/10.48550/arXiv.2504.15088

Let’s Ask Claude!



> *What is Anthropic's definition of "Red teaming" in the context of AI system validation and verification?*

Anthropic's definition of red teaming (from *"Red Teaming Language Models to Reduce Harms: Methods, Scaling Behaviors, and Lessons Learned"*, Aug 22, 2022):

> *Red teaming is the practice of adversarially probing AI systems to* ***simultaneously discover, measure, and attempt to reduce*** *potentially harmful outputs.*

In their framing, red teaming is not purely a testing activity — it is a continuous loop: **discovery** (find what harms the model can produce), **measurement** (quantify how often and how severely), and **reduction** (use those findings to improve the model). Their dataset of 38,961 red team attacks covered a spectrum from "offensive language to more subtly harmful non-violent unethical outputs."

Key operational characteristics from Anthropic's methodology:
- Human red teamers craft adversarial prompts ("attacks") designed to elicit policy-violating or harmful outputs
- Results are analyzed across model types and scales to understand how training (RLHF, rejection sampling, etc.) affects susceptibility
- The goal is community-level norm-setting, not just internal QA — hence their data release

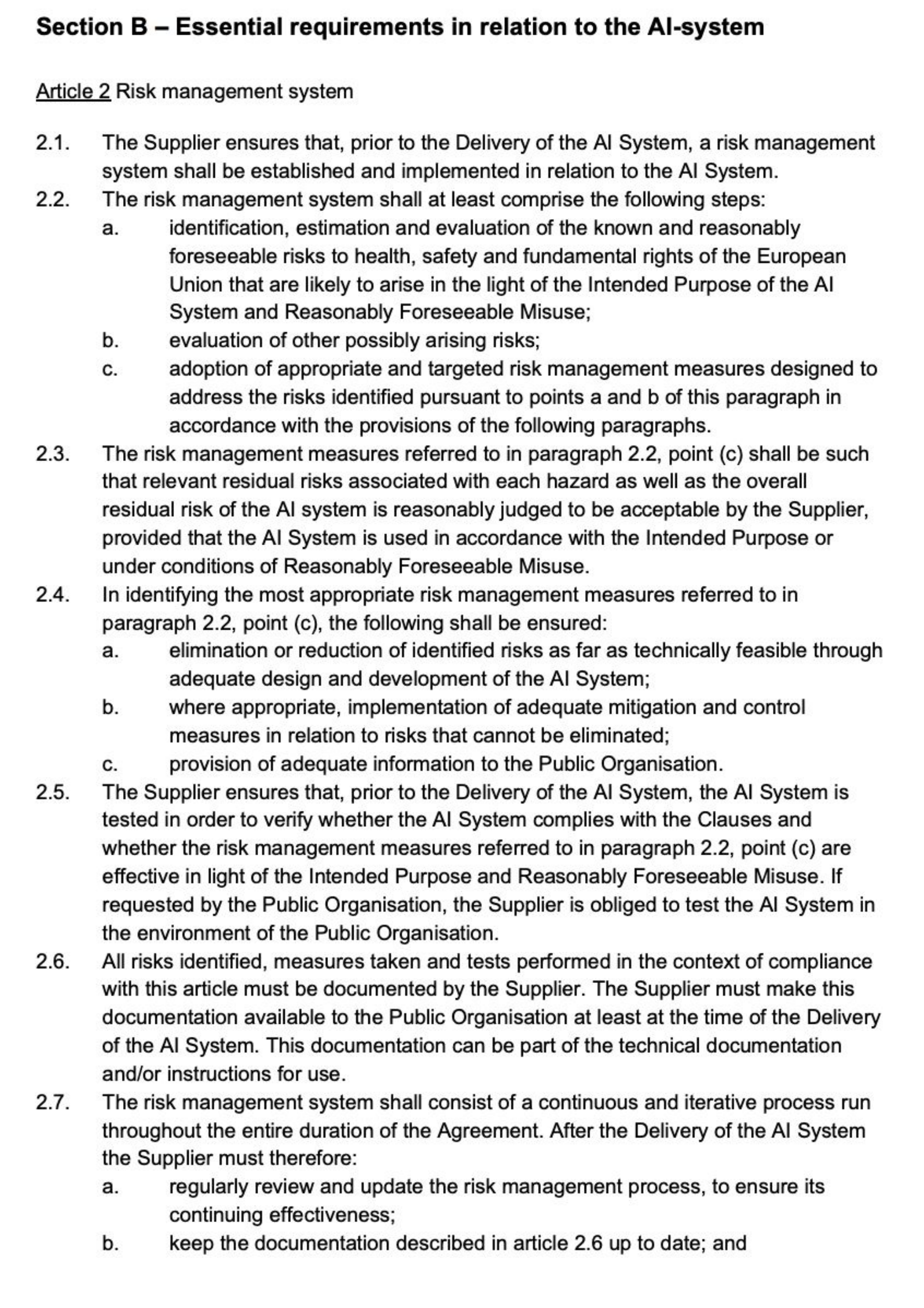
Their https://www-cdn.anthropic.com/82564d4ec2451b2eed2e0796b7c658fc989f0c1a/Anthropic_RedTeaming.pdf expands on the governance framing if you want the policy-level definition rather than the research-level one.

✻ Cooked for 1m 39s

Here’s Where This Gets Dangerous

Look at your AI or Sample AI policy clauses.

If they’re like the templates lots of firms use (and their lawyers are cool with), this Newspeak makes it easy for the AI firms to attest to these things.



Section B – Essential requirements in relation to the AI-system

Article 2 Risk management system

2.1. The Supplier ensures that, prior to the Delivery of the AI System, a risk management system shall be established and implemented in relation to the AI System.

2.2. The risk management system shall at least comprise the following steps:

a. identification, estimation and evaluation of the known and reasonably foreseeable risks to health, safety and fundamental rights of the European Union that are likely to arise in the light of the Intended Purpose of the AI System and Reasonably Foreseeable Misuse;

b. evaluation of other possibly arising risks;

c. adoption of appropriate and targeted risk management measures designed to address the risks identified pursuant to points a and b of this paragraph in accordance with the provisions of the following paragraphs.

2.3. The risk management measures referred to in paragraph 2.2, point (c) shall be such that relevant residual risks associated with each hazard as well as the overall residual risk of the AI system is reasonably judged to be acceptable by the Supplier, provided that the AI System is used in accordance with the Intended Purpose or under conditions of Reasonably Foreseeable Misuse.

2.4. In identifying the most appropriate risk management measures referred to in paragraph 2.2, point (c), the following shall be ensured:

a. elimination or reduction of identified risks as far as technically feasible through adequate design and development of the AI System;

b. where appropriate, implementation of adequate mitigation and control measures in relation to risks that cannot be eliminated;

c. provision of adequate information to the Public Organisation.

2.5. The Supplier ensures that, prior to the Delivery of the AI System, the AI System is tested in order to verify whether the AI System complies with the Clauses and whether the risk management measures referred to in paragraph 2.2, point (c) are effective in light of the Intended Purpose and Reasonably Foreseeable Misuse. If requested by the Public Organisation, the Supplier is obliged to test the AI System in the environment of the Public Organisation.

2.6. All risks identified, measures taken and tests performed in the context of compliance with this article must be documented by the Supplier. The Supplier must make this documentation available to the Public Organisation at least at the time of the Delivery of the AI System. This documentation can be part of the technical documentation and/or instructions for use.

2.7. The risk management system shall consist of a continuous and iterative process run throughout the entire duration of the Agreement. After the Delivery of the AI System the Supplier must therefore:

a. regularly review and update the risk management process, to ensure its continuing effectiveness;

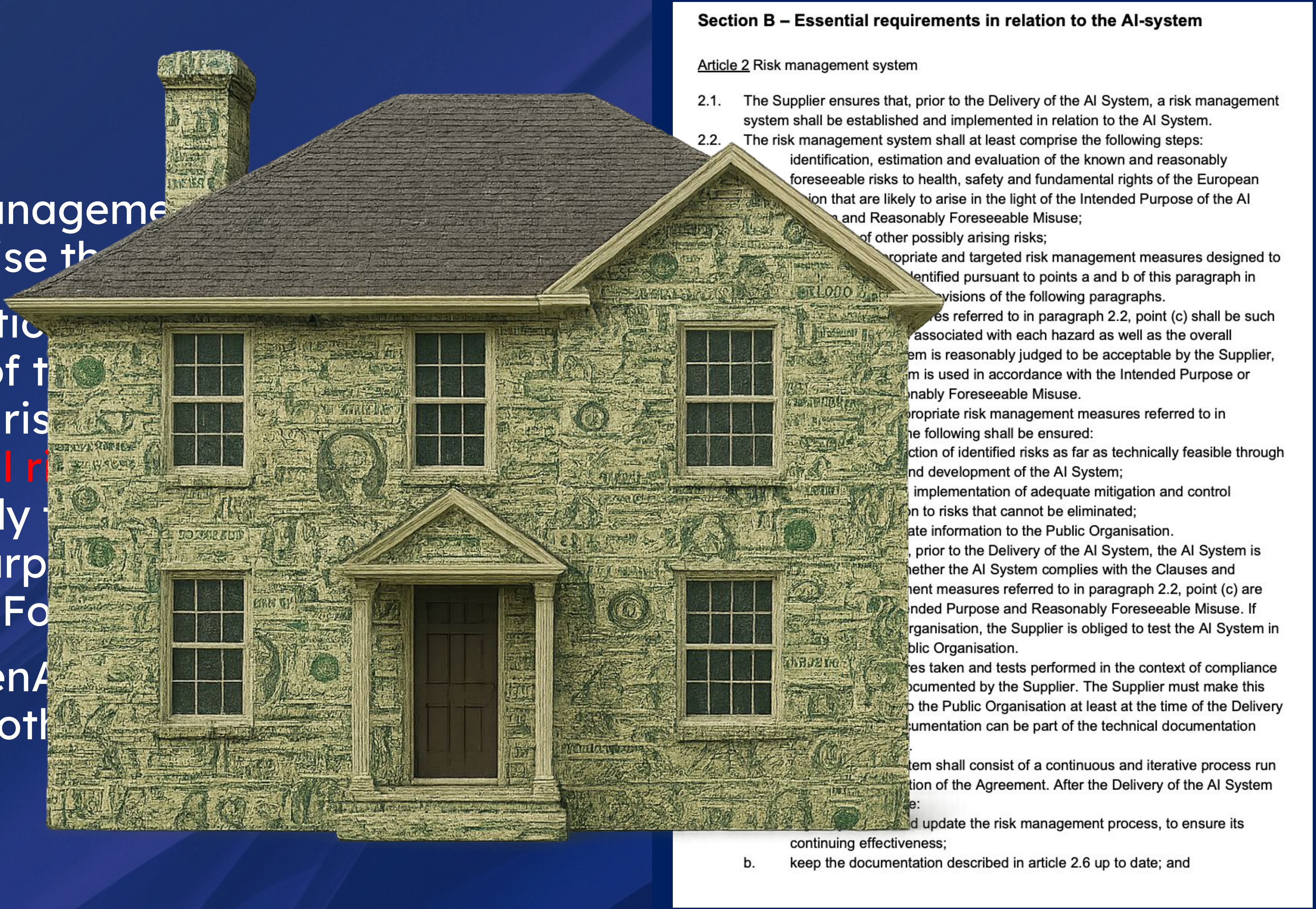
b. keep the documentation described in article 2.6 up to date; and

Like…

“The risk manageme
least comprise th

a. identificatio
evaluation of t
foreseeable ris
fundamental ri
that are likely t
Intended Purp
Reasonably Fo

And the OpenA
pre-form anoth
banknotes.



Section B – Essential requirements in relation to the AI-system

Article 2 Risk management system

2.1. The Supplier ensures that, prior to the Delivery of the AI System, a risk management system shall be established and implemented in relation to the AI System.

2.2. The risk management system shall at least comprise the following steps:

identification, estimation and evaluation of the known and reasonably
foreseeable risks to health, safety and fundamental rights of the European
ion that are likely to arise in the light of the Intended Purpose of the AI
and Reasonably Foreseeable Misuse;
of other possibly arising risks;
ropriate and targeted risk management measures designed to
lentified pursuant to points a and b of this paragraph in
visions of the following paragraphs.
es referred to in paragraph 2.2, point (c) shall be such
associated with each hazard as well as the overall
em is reasonably judged to be acceptable by the Supplier,
m is used in accordance with the Intended Purpose or
nably Foreseeable Misuse.
propriate risk management measures referred to in
he following shall be ensured:
ction of identified risks as far as technically feasible through
nd development of the AI System;
implementation of adequate mitigation and control
n to risks that cannot be eliminated;
ate information to the Public Organisation.
, prior to the Delivery of the AI System, the AI System is
ether the AI System complies with the Clauses and
ent measures referred to in paragraph 2.2, point (c) are
nded Purpose and Reasonably Foreseeable Misuse. If
rganisation, the Supplier is obliged to test the AI System in
blic Organisation.
es taken and tests performed in the context of compliance
ocumented by the Supplier. The Supplier must make this
o the Public Organisation at least at the time of the Delivery
umentation can be part of the technical documentation

tem shall consist of a continuous and iterative process run
tion of the Agreement. After the Delivery of the AI System
e:
d update the risk management process, to ensure its
continuing effectiveness;

b. keep the documentation described in article 2.6 up to date; and

“Make it easy to do it right”

If you’re using “safety” to mean “We won’t Terminator 2 You” and “Red Teaming” to mean “We checked, and our system hardly ever tells you how to whip up a batch of phosphine gas,” it’s hard to know how to do the right thing to lock down endpoints.

Make it easy for people to do the right thing, or they’ll <u>always</u> do the easier, wrong thing.

“We want to make it easy for developers to do the right thing when it comes to infosec.”

— Nicole Forsgren, *Accelerate*

Don’t Trust, and Verify

Assume AI tools come compromised from the factory.

Assume nothing else.

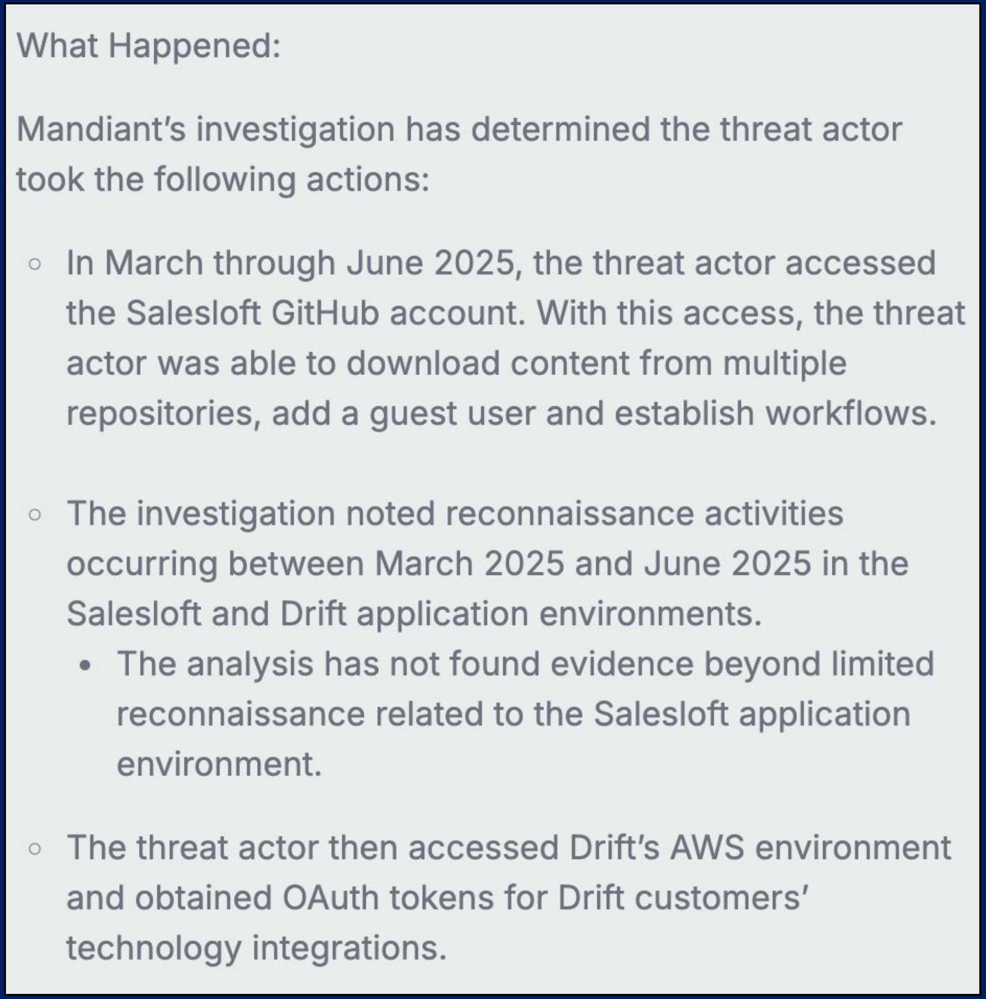
An AI vendor’s SOC2 Type II is as valuable as any other SOC2 Type II.

Presume firms have not crossed their Cyber t’s.

Presume AI firms have not looked at their Cyber t’s.

Disbelieve their words. Ask them what the words mean. Test their answers. Especially when they use words that sound like your words.

Check everything.



What Happened:

Mandiant’s investigation has determined the threat actor took the following actions:

- In March through June 2025, the threat actor accessed the Salesloft GitHub account. With this access, the threat actor was able to download content from multiple repositories, add a guest user and establish workflows.
- The investigation noted reconnaissance activities occurring between March 2025 and June 2025 in the Salesloft and Drift application environments.
  - The analysis has not found evidence beyond limited reconnaissance related to the Salesloft application environment.
- The threat actor then accessed Drift’s AWS environment and obtained OAuth tokens for Drift customers’ technology integrations.

Some Obvious Starting Points

You will find Legal, Audit, IT, and Operations people to be great allies; they're as confused as we are.

Create, review, or update all AI- related policies, procedures, runbooks, and strategies, optimizing for clarity, defining terms.

Like AI itself, teams wanting AI will try anything to get the result they desire.

Create centers of excellence around good AI practices.

Threat model, code audit, lockdown, use monitoring tools.

Like love, AI will find a way (to ingest all your data).

Certified Pre-Pwned: Assume compromise before you tear off the shrink-wrap.

02

“Modeling, not narratives.”

Modeling <u>Your</u> AI Use

Threat Models

Risk Maps

Plain English Breakdowns

Speak About AI Tools In Plain, Easy-to-Understand Terms

Often, C*Os want AI, but can’t express why.

Why do we need this? → What business problem is this tool addressing? How have you tried to address it in the past? Is there a reasonable, sane reason to believe AI can solve it? Why do you believe there is? (Has it been modeled?)

What can go wrong? → Have you drawn a network schematic of how this tool integrates, from where it obtains your data, and where (you believe) it is sending it? Have you table-topped a situation where it is a runaway train? What’s the plan to tackle that?

What training are you conducting? → Have you got a plan to train all the users, plus all the administrators, in the nuances and details of how this thing works?

A Massive Caveat

Check the Sell-By Date Of Advice You Follow; Don’t Base Next Year’s Decision on This Presentation

Things are all a-changin’ →

This year alone we’ve seen a turning of the corner in AI capability; companies that used to depend on harnesses to enforce outcomes have begun ripping them out. Long context will replace RAG for some use cases. ***LLMs and AI agents are becoming ever more capable but the requirements here won’t change***. Measure everything you can.

AI needs to be “set up for success” →

People get job descriptions, goals and KPIs. They need to be “set up for success”. AI agents also need that, which means ***figuring out how to use them effectively***. Just as an employee exit can be a set up failure, you need to put in the work to set up and use AI agents properly, even as things change so quickly.

Measurement Illusion is a real issue →

You want to make sure that you’re measuring, but also that the things you measure surface problems. The question is, “How do we know if our AI investment has paid off - and how can we see if it’s going off-track?” ***Measurement Illusion is when you measure things that feel good but don’t discover problems***.

Modeling, not Narratives

AI Adoption is vastly outpacing Governance. That which is measured may be managed.

“Cost Reduction” → Baseline the target process: headcount, vendor cost, error and rework rate, cycle time. Model what changes when AI handles a defined share of that workload, at what cost, and against what quality threshold. Include the cost of AI tooling. Produce an ROI timeline tied to actual line items, not gee-whiz bullshit.

Careful, though! → Cost reduction is more readily quantified in engineering, where outputs are discrete and volume-trackable. For non-engineering functions like marketing and sales prose creation, run a time-bound survey before deployment: over 10 business days, teams track time on defined tasks. That baseline is what post-adoption measurement is compared against.

Modeling, not Narratives

AI Adoption is vastly outpacing Governance. That which is measured may be managed.

“Increased Productivity” → Task inventory the target function: weekly hours per task category, percentage spent on low-judgment repeatable work, and current output volume. Analyze whether non-deterministic output will produce measurable negatives such as help-desk calls or do-overs. Define a measurable output unit, establish a baseline, and project the new rate under AI assistance. Productivity gain is a ratio, not a narrative.

“Increased Efficiency” → For a targeted workflow, use a current process map depicting current steps, handoffs, wait times, and decision points. Enumerate which steps AI can compress or eliminate and show the resulting cycle time, and include the acceptable error rate at each stage. Efficiency gain is measured against cycle time and handoff reduction, not satisfaction or reduced-effort stories.

03

Blast Radius

Oops.

The Infosec Basics FOMO-Driven CEOs May Forget

Treat AI in your environment like a Chevy Corvair: Unsafe At Any Speed

What am I integrating? → What is the tool? What does it do? Do we need that? What permissions is it [requesting|demanding]? Who makes it? What are the T&Cs? How do they secure their environment? Basic 3PR assessment questions.

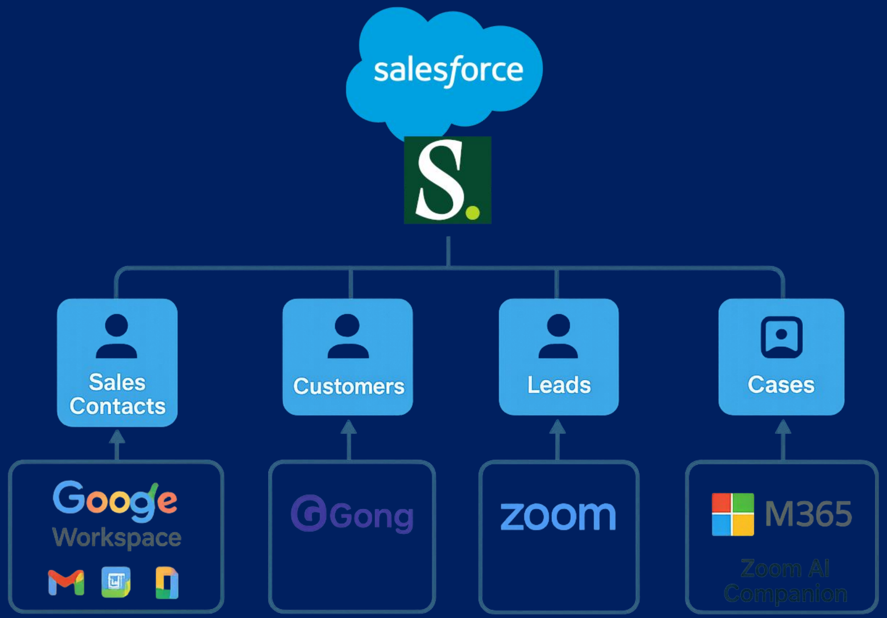
What can go wrong? → Have you threat modeled? Have you risk mapped? Have you sat with a cross-functional group to list not just the benefits, but also the hazards? (All too often I see one or the other.) Monitoring its data-flows? Data blast-radius enumeration?

What training are you conducting? → [We'd hope] prompt engineering? Safety and security training? Incident awareness and recognition? (How we know if something's wrong.) Regular audits of payloads?

What’s The Blast Radius?

Everything you’ve connected it to.

- Assume compromise on day one
- View all your data as potentially ingested, shared by your new AI pals. (And you know what data is in these systems, right? Right?)
    - Is ‘their” AI theirs, or is it a license for someone else’s?
        - What’s the license?
            - Are you sure?
        - Where are they storing your data?
            - Is it Virginia?
                - How do we know?
- Practice fundamental cybersecurity housekeeping

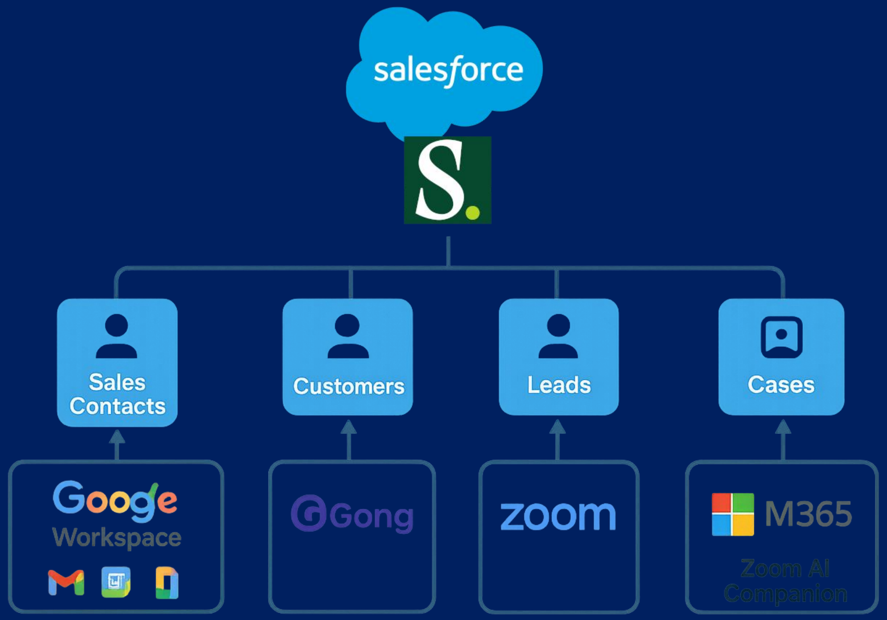


$57.9Bn over 11 funding rounds and ChatGPT couldn’t even draw this right.

Monitor, Monitor, Block; ADR

Everything you’ve connected it to.

- If you’re counting on your SFDC cloud config to limit data, make sure your configuration is what you think, with SaaS Posture & Security Management
    - Make sure your SF admins really understand SF. Make sure they audit randomly, regularly.
- Have you classified your SF data?
- Have you denied the AI tool all but needed data?
- Have you deployed DLP? Is it in block mode?
- Check cloud services (GW, M365). Deny all but required (e.g., calendar, not mail or docs) if possible.
- Ask lots of questions. Disbelieve all the answers.
- Write down the decision and the context in an ADR, so later you’ll remember why



$57.9Bn over 11 funding rounds and ChatGPT couldn’t even draw this right.

04

You Don’t Got This

Get training.

Get Outside Training & Guides to Securing AI

These change regularly. There are many, many, many. Consider this dated. Find your own.

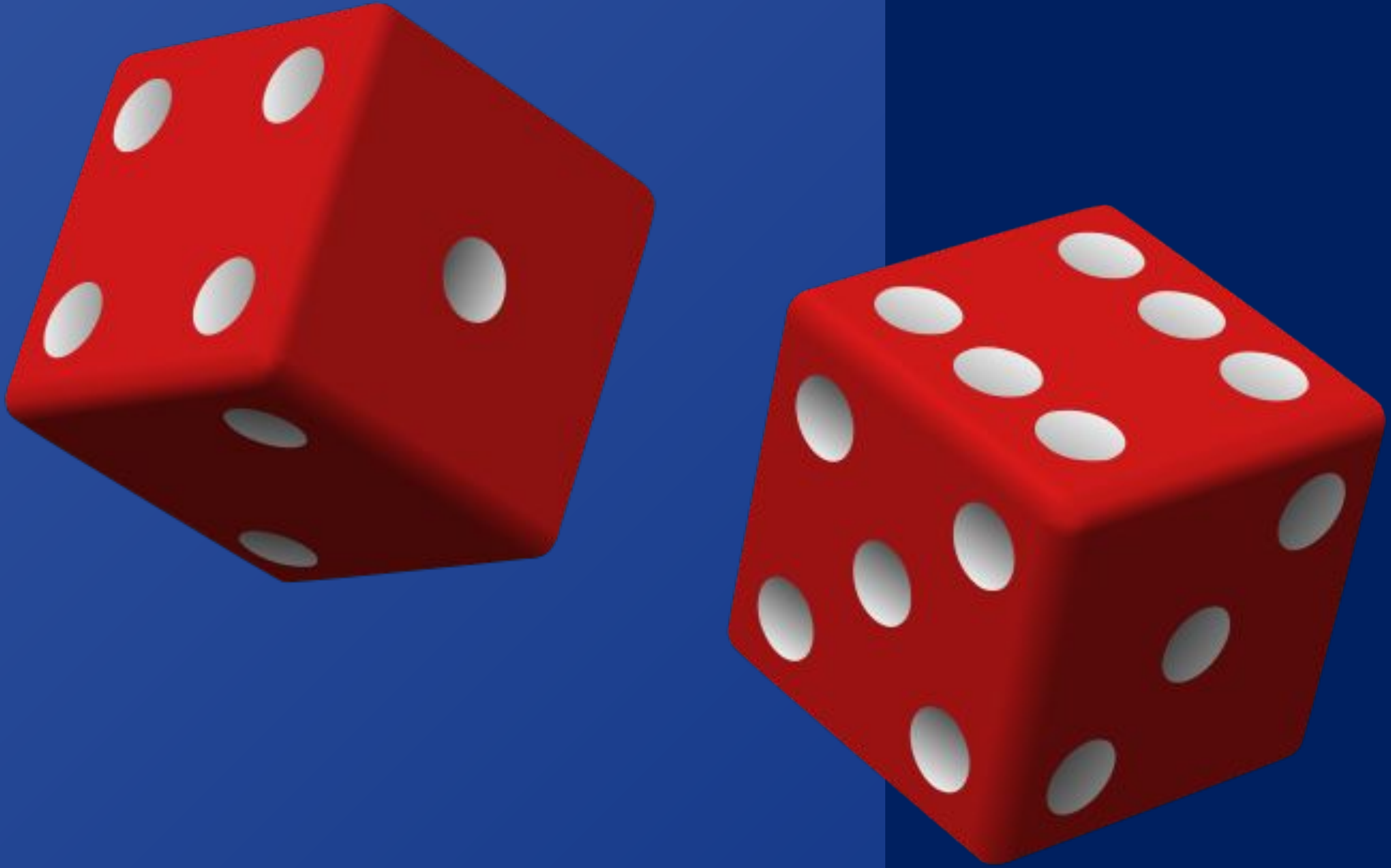
OASIS/Coalition for Secure AI (CoSAI)	→	https://www.coalitionforsecureai.org/
MITRE	→	https://atlas.mitre.org/matrices/ATLAS
Microsoft/OpenAI Anthropic	→	https://www.microsoft.com/en-us/security/security-insider/emerging-trends/ai-security-guide https://chatgpt.com/admin/api-reference https://docs.anthropic.com/en/docs/build-with-claude/prompt-engineering/overview
OWASP AI Exchange “The world’s AI security guide”	→	https://owaspai.org/

Get Outside Training & Guides to Securing AI

These change regularly. There are many, many, many. Consider this dated. Find your own.

Prompt Engineering	→	https://docs.anthropic.com/en/docs/build-with-claude/prompt-engineering/claude-4-best-practices https://learnprompting.org/docs/introduction
Model Context Protocol	→	https://modelcontextprotocol.io/specification/draft/basic/security_best_practices
Agentic Blah Blah Blah	→	https://docs.aws.amazon.com/prescriptive-guidance/latest/strategy-operationalizing-agentic-ai/introduction.html https://platform.claude.com/docs/en/agent-sdk/secure-deployment
Khlaff and Myers West on safety revisionism	→	https://doi.org/10.48550/arXiv.2504.15088

Or, You Can Roll The Dice



Next Steps

Questions?

A PDF of this presentation is available.

Velocity's Edge Podcast on ADRs with Chris Swan and Halvar Flake: epsd.io/blog

Contact Us

EPSD works with senior leaders and boards when execution risk is rising and the right intervention isn't obvious.

Reach out to us directly

https://epsd.io
Signal: fuzztech.01
nick@epsd.io