THE SAFE NORTH

EXECUTIVE ADVISORY BRIEFING // SERIES NO. 01

The 7 Biggest AI Security Mistakes Companies Are Making

AI adoption is accelerating. Security controls aren't. Here is what business leaders must watch in the era of unmonitored deployments.

The True Driver of Risk

AI adoption rarely starts with the security team.

While IT evaluates official enterprise systems, employees independently discover rapid, unapproved workflows.

This organic adoption bypasses conventional oversight instantly:

- An employee signs up for ChatGPT to draft emails.
- A marketing team utilizes a web-based AI writer.
- A developer connects an unapproved IDE assistant.
- A manager uses an unvetted transcriber for meetings.

THE VISIBILITY GAP

None of this adoption goes through standard security approval channels. This is where the risk begins—you cannot govern what you cannot see.

"You cannot govern what you cannot see."

Mistake #01

Assuming 'Approved AI' Is the Only AI Being Used

Many organizations invest heavily in securing officially provisioned AI channels, operating under the false security that these are the only active entry points.

Unmonitored AI channels used daily in workflows:

- Web Chatbots (ChatGPT, Claude, Gemini, Copilot)
- Browser Extensions with AI reading/writing features
- Automated AI Meeting Recorders and Note-takers

THE VISIBILITY MISTAKE

The underlying problem is not the rapid adoption of AI itself. The problem is AI adoption happening without central IT and security visibility.

"Visibility is the prerequisite to security."

Mistake #02

Treating AI Like Any Other SaaS Application

AI tools are not static data silos. Unlike traditional SaaS, they dynamically process, transform, and potentially learn from the information they receive.

Confidential inputs leaving organizational control:

- Proprietary source code and system structures
- Internal financial models and business strategies
- Legal contracts, customer details, and secrets

THE DATA CONTEXT
A modern enterprise needs to define exactly what data employees are allowed to input into AI systems, rather than simply deciding which tools they can open.

"Data is now code. Treat raw text inputs as active instructions."

Mistake #03

Creating an AI Policy Nobody Reads

Drafting a comprehensive 20-page document and storing it in a corporate library is not governance. It provides liability protection, not behavior change.

Five operational questions employees need clear answers to:

- Can I use AI to assist with my daily work?
- Which AI tools are officially approved for use?
- What classification of information can I enter?

THE USABILITY PRINCIPLE
True AI governance must focus on clarity and usability. It should actively make the safe choice the easiest choice for employees.

"Make the secure choice the easiest choice."

Mistake #04

Giving AI Access Without Thinking About Permissions

The next phase of generative AI is not just chatbots answering questions. We are moving toward active, highly connected agentic systems.

AI systems are increasingly being connected directly to:

- Shared file storage, document centers, and email
- Enterprise CRM platforms and customer databases
- Live code repositories and internal APIs

THE ACTION VECTOR

An AI system with access can potentially do far more than generate text—it can dynamically retrieve data and execute actions. This elevates permission structures to a critical security priority.

"An agent with permissions is an active transaction engine."

Mistake #05

Ignoring Prompt Injection

Prompt injection is a fundamental paradigm shift in security. AI systems process both data and system instructions through the same linguistic channel.

Malicious instructions can be hidden inside:

- Incoming emails, customer messages, or support tickets
- External resumes, files, or documents uploaded for review
- Public web pages scraped or searched by the AI system

THE INTERFACE RISK

An attacker does not need to bypass firewalls or write complex code. They simply hide language instructions inside the material the AI naturally processes, dictating how the AI behaves.

"Linguistic inputs are the new untrusted user input."

Mistake #06

Forgetting That AI Creates a New Security Surface

Traditional boundary-focused security paradigms cannot adequately defend dynamic AI integrations. We must transition our questioning:

TRADITIONAL BOUNDARY SECURITY	AI-ERA SECURITY SURFACE
Who has access? Focuses on authenticating specific users and securing endpoints.	**What can the AI see?** Focuses on the data and contextual visibility of the language model.
Is the connection secure? Prioritizes network encryption and perimeter firewalls.	**What instructions can influence it?** Prioritizes identifying and filtering hidden prompt instructions.
What data is stored? Audits static databases and structured tables.	**What data can leave?** Audits real-time data flows, model extraction, and system actions.

"AI introduces an entirely new layer to the corporate security architecture that demands native controls."

Mistake #07

Waiting for AI Adoption to 'Settle Down'

The velocity of artificial intelligence capabilities makes a traditional 'wait and see' governance strategy highly dangerous. AI adoption is not a static trend that will plateau.

The real consequence of waiting:

- AI platforms roll out structural improvements weekly.
- Business teams quickly establish custom habits and daily workflows based around unmonitored systems.
- Re-engineering insecure workflows after they become institutionalized is extremely difficult and highly disruptive.

THE SPEED GAP

Cybersecurity and IT programs must evolve directly alongside employee adoption. Organizations must establish security guardrails today to safely enable scaling tomorrow.

"Work habits solidify faster than policy. Act now."

Executive Action Plan

Start With Visibility First

Before investing in complex security products, executive leaders must first establish absolute clarity. Answer these five foundational questions:

01 **What AI tools are employees already using?**
Conduct audits to discover shadow shadow software bypasses.

02 **What specific data are they putting into them?**
Map classification levels of processed text, files, and code.

03 **Which AI tools are officially approved?**
Draw a bright line between sanctioned enterprise channels and unapproved consumer tools.

04 **What internal systems can AI access?**
Audit active document libraries, databases, and application integrations.

05 **Who actually owns AI risk inside?**
Clarify accountability structures across Security, IT, Legal, and Business teams.

"You cannot govern what you cannot see." — AI Security Advisory Board

AI Security Starts Before the AI Does.

The primary objective of AI governance is not to restrict innovation or halt adoption. The goal is to build secure, transparent foundations so AI can scale safely:

VISIBLE

Complete mapping of active employee workflows and unapproved tools.

CONTROLLED

Strict guidelines on data classes and precise access integrations.

AUDITABLE

Full tracking and active evaluation of inputs, outputs, and system actions.

SAFE TO SCALE

Strategic business confidence to deploy more models, faster and safer.

The companies that win in the era of intelligence won't be the ones that use the least AI. They will be the ones that know how to use it responsibly.