AI Vocabulary Reference • Tier 5a: AI, Privacy & Security

The language of data protection and operational risk • Aurorae Group, LLC

This reference addresses the privacy and security risks that emerge when AI tools enter organizational workflows — risks that are distinct from traditional data security and often invisible until they create real exposure. It is designed for HR leaders, people managers, and operational leaders who handle sensitive data and need to understand what responsible AI use actually requires. It assumes familiarity with foundational AI concepts (Tier 1).

Data and Privacy Fundamentals

The core concepts that govern how data should be handled in AI contexts — and why the stakes are higher than most people realize.

Personally Identifiable Information (PII)	Any data that can be used to identify a specific individual — including names, email addresses, employee IDs, compensation figures, performance ratings, medical accommodation records, and demographic data. In AI contexts, PII includes not just structured data fields but any narrative or conversational content from which an individual could be identified. **Why it matters for leadership:** HR handles more PII than almost any other function in an organization. Every AI prompt that includes employee names, situations, or details is a potential PII exposure event. The question is not whether HR uses sensitive data — it is whether that data is being handled with the same care in AI contexts as it is in traditional HR systems.
Data Minimization	The practice of using only the data necessary to accomplish a specific task — and no more. In AI contexts, data minimization means stripping prompts of unnecessary identifying details, generalizing scenarios before inputting them, and resisting the temptation to provide context that the AI does not need to produce a useful output. **Why it matters for leadership:** Data minimization is one of the most immediately actionable privacy practices for HR teams. Before inputting any scenario or situation into an AI tool, ask: does this system need to know who this person is? In most cases, a generalized or anonymized version of the situation produces an equally useful output with significantly lower risk.
Contextual Integrity	The principle that information flows appropriately when they match the norms of the context in which information was originally shared. Data shared in confidence in an HR conversation carries different norms than data entered into an AI system that may process, store, or use it in ways the original sharer did not anticipate. **Why it matters for leadership:** Contextual integrity is the reason that 'technically allowed' is not the same as 'appropriate.' An employee who shares a personal situation with HR has not consented to that information being processed by a third-party AI system. Leaders who think carefully about contextual integrity make better decisions about what belongs in an AI prompt and what does not.
De-identification	The process of removing or obscuring identifying information from data before it is used or shared — replacing names with roles, specific figures with ranges, and identifying details with generalized descriptions. De-identification reduces but does not eliminate privacy risk.

	Why it matters for leadership: De-identification is a practical first line of defense for HR teams using AI. It is not foolproof — sufficiently detailed de-identified data can sometimes be re-identified — but it meaningfully reduces exposure for most routine use cases. Building de-identification into prompting practice as a habit is more effective than relying on policy alone.
Data Residency & Sovereignty	Data residency refers to the geographic location where data is stored and processed. Data sovereignty refers to the legal jurisdiction governing that data. Cloud-based AI tools may store and process data in multiple countries by default, subjecting it to the laws of those jurisdictions rather than the organization's home country. **Why it matters for leadership:** For organizations operating across jurisdictions or subject to specific data protection regulations, data residency is not a technical detail — it is a compliance requirement. HR data involving employees in the EU, for example, may be subject to GDPR regardless of where the organization is headquartered. Vendor contracts should specify data residency explicitly.

AI-Specific Privacy Risks

The risks that are unique to — or significantly amplified by — AI systems, as distinct from traditional data handling.

Input Privacy Risk	The risk that data entered into an AI system — in prompts, uploaded documents, or integrated data sources — is retained, used for model training, or accessible to the vendor in ways that compromise confidentiality. Input privacy risk varies significantly between free consumer tools and enterprise platforms with data protection agreements. **Why it matters for leadership:** Input privacy risk is the most common and least understood AI privacy risk in organizational settings. Many employees using free versions of AI tools do not realize that their inputs may be used to train the model or reviewed by vendor staff. This is not a theoretical risk — it is documented in the terms of service of most consumer AI platforms.
Output Privacy Risk	The risk that AI-generated outputs inadvertently expose sensitive information — by reproducing details from training data, combining information in ways that reveal confidential content, or generating outputs that contain more identifying detail than the prompt intended. **Why it matters for leadership:** Output privacy risk is less visible than input risk but equally real. An AI system trained on organizational data — or prompted with sufficient context — may surface information in its outputs that was not intended to be disclosed. HR teams should treat AI outputs involving sensitive topics with the same scrutiny they apply to inputs.
Re-identification Risk	The risk that de-identified data can be reverse-engineered to identify individuals — particularly when AI outputs are combined with other available information. A scenario described without names may still be identifiable if it contains enough specific detail about role, location, situation, or timing. **Why it matters for leadership:** Re-identification risk is why de-identification is necessary but not sufficient. In small teams or organizations, a de-identified scenario may still point clearly to one person. Leaders should calibrate the level of de-identification to the specificity of the situation and the size of the population it describes.

Memory and Retention	Whether and how an AI system stores conversation history — across sessions, across users, or accessible to vendor staff. Memory and retention practices vary significantly across platforms and plan types. Some systems retain no data beyond the immediate session; others store conversations indefinitely by default. **Why it matters for leadership:** Most employees do not know whether the AI tool they are using retains their conversations. This matters enormously for HR use cases involving sensitive employee situations, confidential organizational information, or legally privileged communications. Knowing your platform's retention practices is a basic governance requirement — not an advanced one.
Third-Party Data Exposure	The risk that integrations between AI tools and other organizational systems — email, calendar, HRIS, document management — create unintended data pathways, exposing information to the AI system or its vendor beyond what was explicitly intended. **Why it matters for leadership:** As AI tools become more deeply integrated into organizational workflows, the surface area of potential data exposure expands. A plugin that connects an AI tool to your email system may give that tool access to far more organizational data than the specific task requires. Each integration point requires its own privacy review — not just the primary tool.

Security Risks in AI Workflows

The attack surfaces and vulnerabilities that emerge when AI tools are integrated into organizational operations.

Prompt Injection	A security vulnerability in which malicious or unauthorized instructions are embedded in content that an AI system is asked to process — such as a document, email, or web page — causing the AI to execute those instructions as though they came from a trusted source. A cyberattack specific to AI-integrated workflows. **Why it matters for leadership:** As organizations integrate AI into workflows that automatically process external documents, emails, and data, prompt injection becomes a meaningful operational security risk. Content that looks like data to a human may contain instructions to an AI. AI-integrated workflows that touch external or untrusted content require specific security review.
Shadow AI & Compliance Risk	The specific compliance exposure created when employees use unapproved AI tools with organizational data. Shadow AI is not just a governance problem — it is a compliance problem when unapproved tools lack the data protection agreements, security certifications, or audit trails required by organizational policy or regulation. **Why it matters for leadership:** Shadow AI compliance risk is largely invisible until something goes wrong. An employee who uses a free consumer AI tool to process a benefits dispute or draft a termination letter may have created a compliance exposure that HR will be held accountable for — without any awareness that the risk was created. Governance and clear communication are the only mitigations.
Credential & Access Risk	The security risk created when AI tools are granted access to organizational systems — email, calendars, file storage, HRIS — through integrations, plugins, or API connections. Compromised AI tools or vendors can become vectors for unauthorized access to organizational systems and data.

	Why it matters for leadership: Every AI integration that touches organizational systems is a potential attack surface. Leaders authorizing AI tool integrations should apply the same security scrutiny they would to any third-party system access — including reviewing what permissions are requested, what data is accessible, and what happens to access credentials if the vendor relationship ends.
AI-Generated Phishing	The use of generative AI to produce highly personalized, contextually convincing phishing emails, messages, or communications at scale. AI dramatically lowers the barrier to sophisticated social engineering by enabling attackers to craft individualized, grammatically flawless, contextually appropriate messages without the errors that previously made phishing detectable. **Why it matters for leadership:** AI-generated phishing represents a step-change in social engineering risk that most organizational security training has not yet caught up to. HR teams are frequent targets — payroll fraud, benefits data theft, and employee impersonation are common attack vectors. The signals that previously flagged phishing — generic language, grammatical errors, implausible scenarios — are no longer reliable.
Vendor Security Posture	The overall state of a vendor's security practices, certifications, policies, and track record — including SOC 2 compliance, data encryption standards, breach history, incident response practices, and the security provisions in their contractual agreements. **Why it matters for leadership:** Vendor security posture is the aggregated answer to the question: can we trust this organization with our data? It is assessed through a combination of certifications (SOC 2 Type II is a reasonable baseline), contractual review (data processing agreements, breach notification timelines), and direct inquiry. Vendors who resist security review are themselves a security signal.