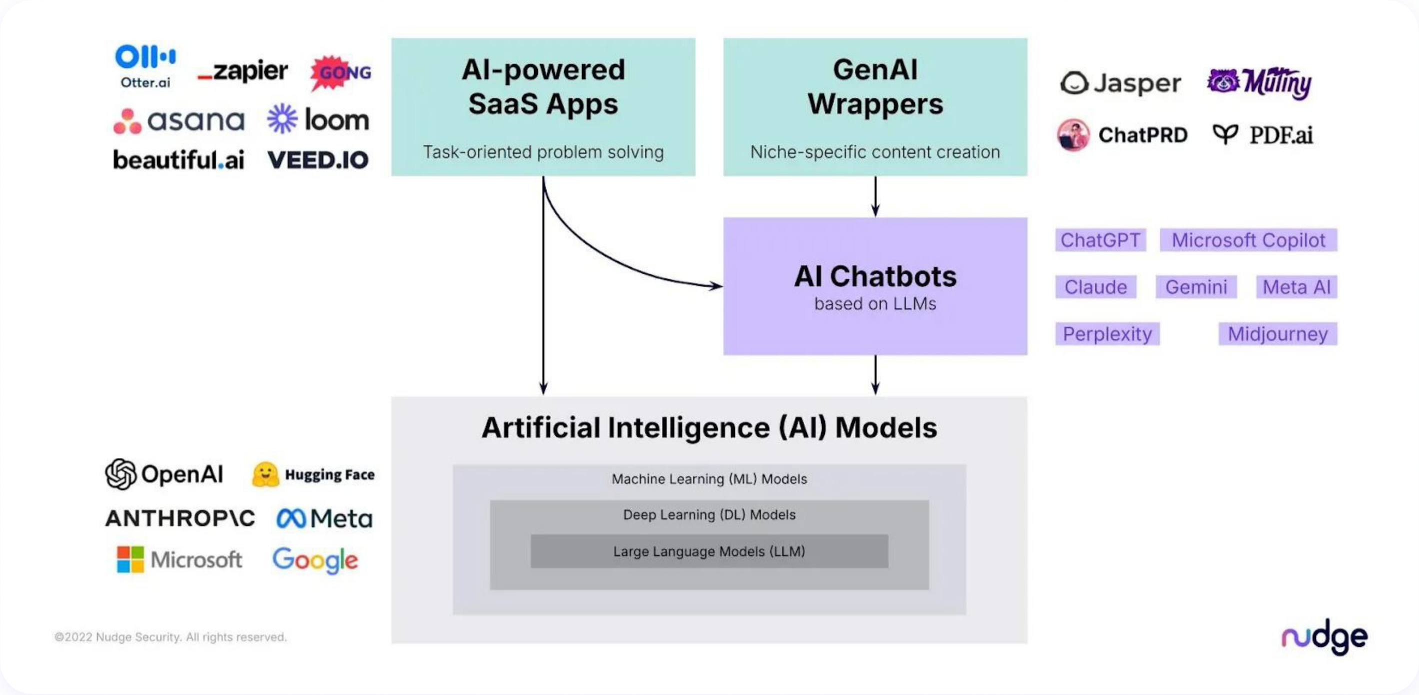
The Practitioner’s Guide to Conducting an AI Risk Assessment

Learn to better equip your organization to ensure safe and compliant adoption of AI tools.

In today's fast-paced technological landscape, AI is no longer a futuristic concept; it's a present-day reality transforming how businesses operate. For security and IT leaders, understanding and managing the risks associated with AI implementation is crucial to safeguarding organizational data and reputation. This guide provides a structured approach to running an AI risk assessment, ensuring safe and compliant adoption of AI tools in your enterprise.

Exploring the AI landscape



The foundation of AI starts with machine learning, deep learning, and large language learning models (LLMs). The most common models include ChatGPT from OpenAI, Claude from Anthropic, and Llama from Meta. Many of these companies have created a chatbot interface with the models; this is what the general public knows as ChatGPT, Claude, and Meta AI, respectively.

GenAI apps don’t stop there. A flood of startups are seizing on the demand for AI by building purpose-built solutions on top of these AI models’ APIs. These GenAI “wrapper” apps aim to reduce the learning curve of prompt engineering with a user-friendly UI designed for specific use cases and outcomes. Since they don’t require a lot of heavy infrastructural development, GenAI wrappers can be launched quickly and easily as a weekend side project, which may suggest that rigorous security controls are not properly in place

Finally, there are “AI-powered” SaaS apps: the multitudes of SaaS providers that want to capitalize on the novelty of AI, boost top line revenue, and stay ahead of the competition by embedding AI-powered capabilities in their offerings. “AI-powered” could mean anything from using one of the common LLMs to surface documentation faster to actively delivering suggestions, results, and value within the product.

The bottom line: the AI landscape is vast and growing exponentially faster. In fact, AI growth trends from Nudge Security show that the number of unique GenAI tools has roughly doubled each quarter starting in 2023. t's critical to keep up with the pace that GenAI tools are created and used by your employees and their SaaS tools.

5 key factors of AI risk assessment

1. Discovery: What AI tools do our employees use?

The first step in any AI risk assessment is identifying all AI-related accounts, users, and applications within your organization. This process involves cataloging not only known GenAI tools in use, but also uncovering new, niche tools that may have slipped under the radar, and any AI-powered SaaS apps. There are five ways to discover what GenAI tools are being used at your organization, providing various levels of visibility.

- **Manual project cataloging:** Teams that want to take an organized, business-aligned approach to exploring and adopting AI often start here. Manual inventory inevitably gets outdated quickly, especially without owners to keep the list updated regularly as business needs change.
- **Network monitoring:** Inspecting traffic details can provide a quick baseline of commonly known GenAI tools, like ChatGPT. These solutions rely on a user-generated database of information to identify specific apps, meaning you have to tell the tool which GenAI apps to look for, rather than the tool discovering new ones for you. While it's a better place to start than a manual inventory, you could still have blind spots when it comes to new, niche GenAI tools—like those that just launched on Product Hunt over the weekend.
- **App integrations:** Analyzing app integrations takes a very narrow view of what GenAI tools are used and connected to core work apps. If you are only concerned with AI risk to critical resources within the organization, this approach may provide a deep view into the connected ecosystem of apps your workforce uses. The downside to this approach, though, is that it is a manual process and may miss any experimentation happening outside of GenAI apps deemed business critical.
- **Agents (e.g. browser extensions, endpoint protection):** While agents can help uncover when and where GenAI tools are used, they may be limited in their visibility of net-new GenAI apps. They also collect way more data than you can put into action. With this approach, it's especially important to be mindful of privacy concerns and ensure proper data handling procedures are in place.
- **Machine-generated email discovery:** Every new account sign-up typically triggers a machine-generated email sent to the email address associated with the new account. Taking an identity-first approach, email discovery covers many of the gaps the other solutions pose from automatically recognizing new, niche GenAI tools to discovering AI use even when the person authenticates with simply a username and password.

2. Trust: How do AI vendors handle security and privacy?

A recent study found that 61% of companies have been impacted by a third-party breach. Given the pace at which GenAI tools have entered the market (many without security programs), it's vital to determine the security posture of GenAI tools and ensure they align with your organization's security standards. (This is especially concerning when 90% of the 2,500+ GenAI vendors Nudge Security has automatically discovered and catalogued have fewer than 50 employees.)

When reviewing GenAI vendors, here are some questions to consider:

- What security certifications does the vendor hold?
- How mature is their security program?
- How do they manage data privacy?
- Have they experienced any recent data breaches?

While security questionnaires can cover some of these questions, conducting these reviews can be time intensive and impede workforce productivity if it is a requirement before using every tool. It can be helpful to steer employees towards already vetted and approved GenAI tools rather than continuing a never-ending stream of GenAI vendor security reviews.

Note: Nudge Security provides free, publicly available security profiles for thousands of SaaS tools, including an expanding list of GenAI tools.

3. Integrations: What GenAI tools are directly connected to core business apps?

GenAI tools often connect to other systems within your organization, creating points where data leaks could happen if not properly managed. A detailed integrations review helps map out these connections and assess their security implications. Key considerations include:

- What systems are connected to the GenAI tools?
- Are they connected via API or OAuth grant?
- What are the scopes for OAuth grants used to connect GenAI tools to other apps?
- What data is being shared with these connected GenAI tools?
- How are the identities managing the integrations protected?
- What guardrails are in place to prevent data from being leaked?

4. Supply chain: What SaaS apps use AI under the hood?

According to recent SecurityScorecard research, 75% of third-party breaches targeted the software and technology supply chain. SaaS vendors have been launching AI-powered functionality at an accelerated pace since ChatGPT went viral in December 2022, so it is critical for third-party risk management teams to stay on top of which vendors are adding AI to their sub-processor list and supply chain. To manage third and fourth party risk, this review should regularly investigate:

- What new AI has been added to the software supply chain or sub-processor list of their third-party vendors?
- Do these new GenAI tools align with my company's risk appetite for AI?
- How are these SaaS vendors handling data privacy with the new additions of AI in their pipeline?
- How do these SaaS vendors intend to respond to any data breaches that their GenAI tools experience?

5. Communication: How do users know what safe, acceptable AI use looks like?

Employees want more AI education. In fact, in a recent EY survey, 81% say they would feel more comfortable about using AI if best practices on responsible AI were routinely shared. IT and security leaders have an opportunity and responsibility to ensure that employees are aware of the organization's acceptable use policy and AI best practices. Regular training sessions, clear communication channels, and accessible support resources can help reinforce these policies.

Ask yourself:

- Have all users been trained on the acceptable use policy?
- What automation can I put in place to ensure my employees are educated on our AI policy as they begin to use GenAI tools?
- Do they know where to find help if they have questions or encounter issues?
- Do they know what approved GenAI tools are available to them?
- Do they know how to request approval for a new GenAI tool or new project involving AI?

How Nudge Security can help

By embracing a structured approach to AI risk assessment, leaders can not only safeguard their data and reputation but also unlock AI's transformative potential securely. Encouraging a culture of vigilance and continuous improvement positions your organization at the forefront of innovation while maintaining robust security protocols.

Explore how Nudge Security can help you conduct your GenAI risk assessment and streamline your on-going genAI security and governance efforts. Learn more →

Get started with a 14-day free trial with zero commitment.
nudgesecurity.com