CONCENTRIC AI

GenAI: A Generational Paradigm Shift

Redefining Data Security in a New Era

Authors:

Dr. Madhu Shashanka, Chief Scientist and Co-Founder, Concentric AI

Lane Sullivan, SVP, Chief Information Security and Strategy Officer, Concentric AI

Table of Contents

Introduction 4

What Exactly Is GenAI, and Why Is It a Big Deal? 6

- Software Eats the World 8
- Machine Learning: The Almost Forgotten Hero 8
- Why Is GenAI Different? 9
    1. Representation Learning – Goodbye Feature Engineering! 10
    2. General-Purpose Models 11
    3. Natural Language "Instruction Following" 13
- What's the Catch? 14
- Intelligence and Automation 15
- Autonomy and Predictability: The Path to "Agentic AI" 16
- Are We Close to Artificial General Intelligence, or AGI? 18

GenAI Impact 19

- Trading Off Expertise and Operational Complexity for Computing 19
- Democratizing Intelligent Automation 20
- Innovation Velocity 21
- Is GenAI Going to Replace Technology Workers? 21

Bottom Line: A New Layer of Enterprise Infrastructure 22

- A Familiar Story: The BYOD Parallel 23

Collateral Repercussions 24

- Upending the Principle of Least Privilege 25
- The Expanding Risk Surface 26
  - Input Leakage 26
  - Output Exposure 26
  - Accessibility Without Oversight 27
  - Second-Tier Supply Chain Risk 27
  - Governance Gaps in Training Data 27
  - Application Code Risk 27
- Data Security Risk Governance 29
- Why the Cost of Inaction Is Higher 30

The Path Forward: Context-Driven Visibility and Layered Controls 31

- Inside-out Governance (Data Hygiene) 31
- Outside-In Controls (GenAI Monitoring) 32

The Regulatory Horizon 33

Looking Ahead 34

Introduction

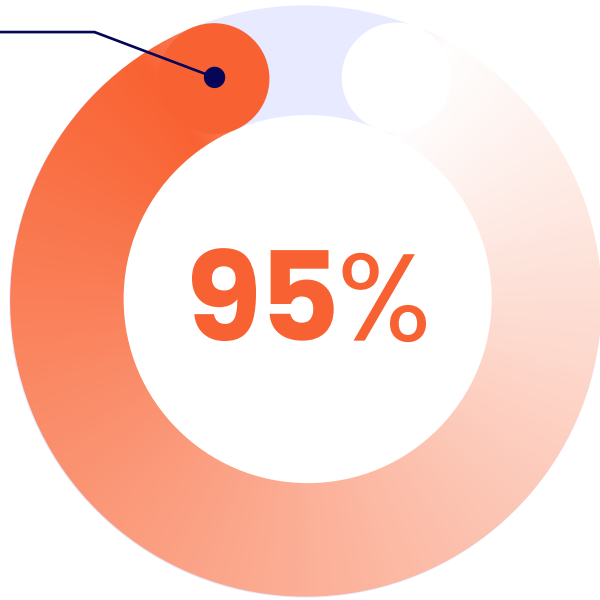
Generative AI (GenAI) catapulted from a curiosity to a central force in enterprise technology almost overnight. Its ability to generate text, code, images, and insights on demand has made it indispensable for employees eager to cut through complexity and accelerate productivity. But with this innovation and efficiency comes massive exposure to risk.

In calls with executives and AI governance leaders across industries, one theme surfaces again and again: data security has moved from a key concern to the focal point of their strategy, and it is the defining challenge of enterprise AI adoption. Unlike traditional software or even past waves of machine learning, GenAI changes the fundamental process for securing data in an organization.

Recently, we were pulled into a call with a prospect—a large enterprise reviewing Concentric AI as a potential vendor to address some of their data security needs. The call was part of their vendor risk management process, and we were supporting the efforts of our champion on their data security team to get approval from their internal AI governance team. What started off as an operational call quickly turned into a discussion about the risks they were facing due to GenAI, how it was exponentially exacerbating data security risks, and how Concentric AI could help them address some of those challenges.

This is just one example of several such conversations we have been having—with CIOs, CISOs, customers, partners, and security practitioners at large. GenAI is new and exciting, but it also comes with risks—especially around data security—and enterprises are grappling with a rapidly evolving landscape.

A recent MIT Study claims that 95% of enterprise GenAI pilots are failing.



There must be a reason why, and it's not because the technology is weak. Rather, it's because enterprises lack the governance and security frameworks to operationalize GenAI appropriately and responsibly. In another study, also from MIT, enterprise leaders cited data security as the top business risk and security risk hindering faster AI adoption. Meanwhile, CSO Online has flagged "shadow AI," which is unsanctioned employee use of public tools, as a driver of skyrocketing data risks beyond corporate control.

The stakes are clear: GenAI is not a bolt-on productivity tool; it is a new infrastructure layer that massively expands the enterprise data surface. For boards and CEOs, this is not a technology issue, it is a business risk, a fiduciary obligation, and a competitive differentiator.

CISOs cannot treat GenAI as another DLP program, it is technology that requires reframing the entire security operating model.

The enormous amount of hype and noise around GenAI has led to confusion, uncertainty, and a fear of missing out. This white paper attempts to bring clarity by sharing Concentric AI's perspective on what makes GenAI so different, what the impact is—both intended and unintended, and what business leaders can do to minimize risk and enable innovation.

What Exactly Is GenAI, and Why Is It a Big Deal?

GenAI entered the popular parlance with the release of ChatGPT almost three years ago now. The "generative" descriptor for AI comes from the technical distinctions between "generative models" that underlie this approach to AI as opposed to "discriminative models" that are heavily used in applications of machine learning (ML).

Without going into technical details, what this means is that generative models can be used to generate new instances of data. This ability to generate new and novel content—whether it is text with applications like ChatGPT, or whether it is other forms of content such as audio and video from models like Sora—is what distinguishes "GenAI" for the average person.

Gen AI might seem mostly like a novel toy for the average person, but it already has had an enormous and undeniable impact on productivity and efficiency across a variety of business use cases. And we have barely scratched the surface in terms of innovation in this rapidly evolving field.

However, we are starting to see tempered expectations this year from businesses. There have been reports of disappointing results, unrealized payoffs, and frustration despite relatively high adoption.

Gartner, in its latest Hype Cycle for AI, has placed GenAI in the "Trough of Disillusionment."

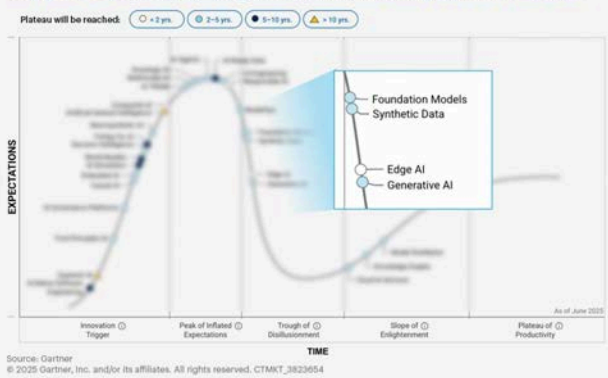


Figure 1: Gartner Hype Cycle for Artificial Intelligence, 2025

But people tend to overestimate the impact of technological advances in the short term while vastly *underestimating* the impact over time. While GenAI, having already broken into the popular imagination, is going through a reality check from the initial hype, we strongly believe it will lead to a generational paradigm shift.

There are three main factors defining the GenAI era that are driving this change:

- **Representation learning**
- **General-purpose models**
- **Natural language "instruction following"**

Before diving into these, it will help to review how we got here. Business impact and transformational changes are primarily driven by automation. At a high level, this process starts with business goals (the "why"), which are then transformed into operational and technical requirements (the "what"), before being followed by execution (the "how").

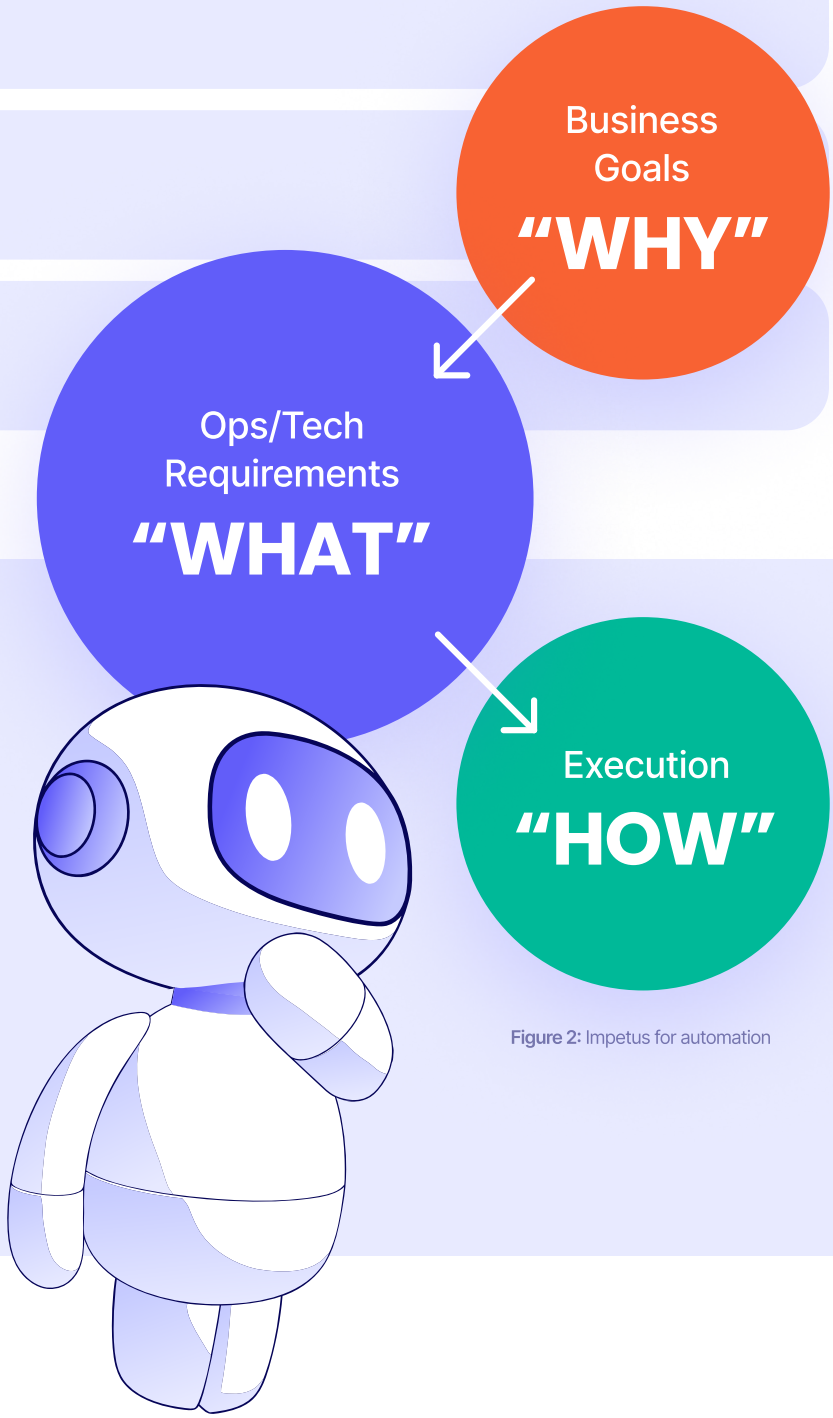


Figure 2: Impetus for automation

Software Eats the World

When Marc Andreessen famously said back in 2011 that "software was eating the world," he was emphasizing how every business, regardless of the industry or vertical it operated in, was ripe for transformation because of software automation. Software is powerful because it allows effective and reliable automation at scale.

Figure 3 summarizes the software automation framework. Given a task (the "what"), one must define details of the approach (the "how") and then execute to develop the software. The software, once deployed, achieves desired automation.

Figure 3: The software automation framework

Machine Learning: The Almost Forgotten Hero

As most people are aware, artificial intelligence (AI) is not new. While every problem today looks like a nail for the proverbial GenAI hammer, the AI field as an academic area of research has existed for several decades now. Before GenAI hype took over, AI had already achieved massive impact across industries via machine learning, which is a sub-field of AI.

The need arose in the early days of software engineering where, for certain kinds of problems, even domain experts found it incredibly hard—and in certain cases impossible—to accurately specify and define all the steps necessary to achieve a task.

One of the best examples to illustrate this is the task of "chicken sexing" where experts can effortlessly tell apart the genders of newly hatched chicks but are at a loss to explain how they do it. For more on this, read James McWilliams' fascinating overview of the profession and its history. However, if there is enough data of inputs and desired outputs, statistical algorithms can be applied to "infer" how the transformation happens. And one can apply this knowledge, in the form of a "machine-learning model," on new inputs to transform them into desired outputs.

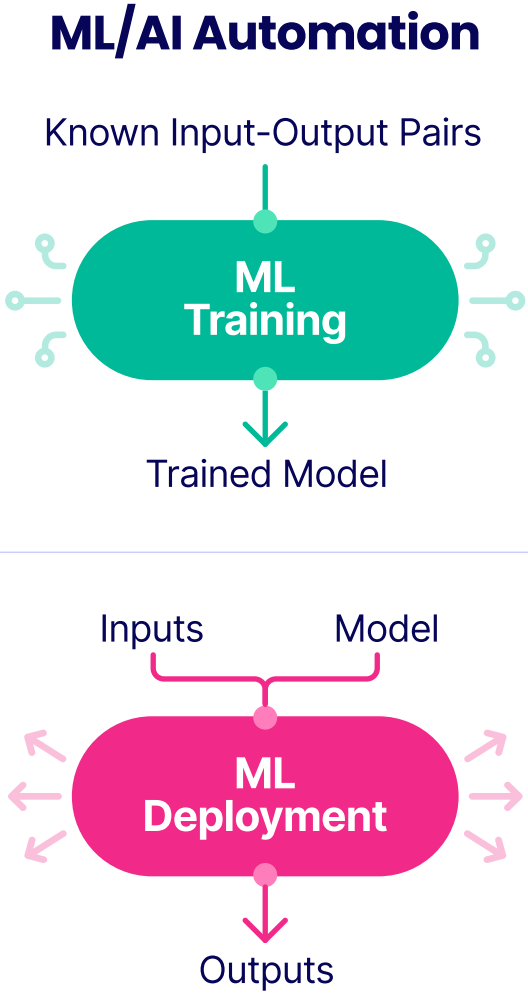


Figure 4: Knowledge applied to new inputs in the form of ML creates desired outputs

While the end goal of ML systems is automation as well, they exhibit "intelligence" unlike traditional software by "learning" the implicit relationships between input-output pairs in the data—without being instructed by domain experts, and they can apply that knowledge to achieve automation.

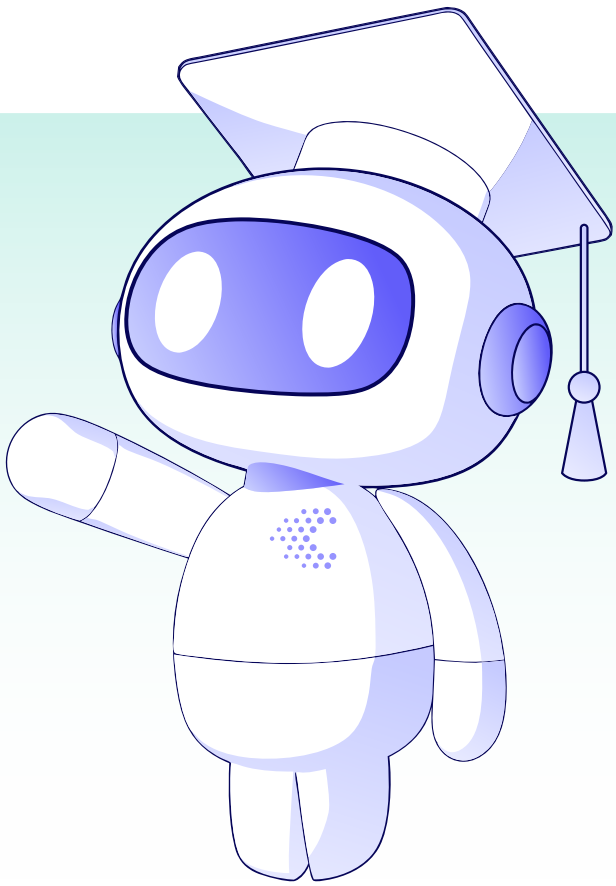
Why Is GenAI Different?

Coming back to the world of GenAI today and the three main factors distinguishing it from the types of AI and ML from the previous era. And we have rapid advances in deep learning, starting in the early and mid-aughts, to thank for that.

1

Representation Learning – Goodbye Feature Engineering!

In traditional machine learning, human expertise plays an important role in the development of models. Specifically, domain experts with knowledge of the task to be automated need to come together with machine learning experts to "represent" and "code" the inputs the right way so that developed models can achieve automation with high accuracy.



For example, consider the task of detecting spam emails. One can take a dataset of several known spam and legitimate messages and train a machine-learning model from this labeled dataset. But what gets fed as input into machine learning is not the data itself. The data first needs to be transformed into a set of "features," which are then fed into the machine learning system.

For detecting spam, example features could be the number of misspelt words, ratio of numeric characters to alphabetic characters, proportion of non-alphanumeric characters such as punctuations and emojis, and so on. The features to be used are often decided by domain experts based on their knowledge and intuition. This is an important step in traditional machine-learning approaches and is often termed "feature engineering."

The performance of the machine-learning model depends on and is also limited by the appropriateness and quality of the features chosen to represent the data. This can be a double-edged sword: it can work well where designers have deep expertise about the application task, but it can also be very limiting if designers do not have a good intuition for what makes high-quality features.

Figure 5: Unlike ML, deep-learning models permit "representation" learning

This is where deep learning has changed the game. Instead of first transforming the data into a set of features, you can feed the data—a string of words and sentences within these files—directly into the deep learning training system and it will internally "learn" the most appropriate set of "features" along with the machine-learning model. Instead of having to hand-code the right set of features, deep learning allows "representation learning" as well so that the data is represented by the "best" set of features.

This paper will not get into the technical details of deep learning, but interested readers can refer to Concentric AI's ebook for an overview of key concepts.

2

General-Purpose Models

While machine learning focuses on building task-specific models, GenAI models focus on characterizing the distributions of your input data, independent of the eventual task for which the model might be used. In other words, one can create GenAI models without a specific task in mind ahead of time.

This started with large language models (LLMs), trained on large volumes of text data, that could encode semantics of language without experts having to do any "feature engineering." The current generation of LLMs are trained on "humanity-scale" text datasets that incorporate significant amounts of textual data that humanity has generated to date. Since much of the text that we have generated represents our understanding of the world, the models thus trained also start reflecting an understanding of our world—beyond text and language—having some implicit representation of the world under the hood.

These models are now more generally referred to as "foundation models" and encompass models trained on other data modalities such as audio, images, video, and so on. While these models can be further refined, or "fine-tuned," for specific tasks, the "base models" without any fine-tuning can be used for a variety of tasks.

In the early days of GenAI in 2020/2021, there was a robust debate on its limits and whether the models were capable only of regurgitating patterns seen in the data, essentially being "stochastic parrots," or whether they could truly "extrapolate" beyond the situations seen in training.

There had been early clues predating this debate that reinforcement learning could make these models more than statistical interpolators and not just seem intelligent due to their sheer sizes and scale.

Back in 2016, there was a series of Go matches arranged between Lee Se-dol, the reigning world champion, and AlphaGo, a deep-learning model developed by DeepMind. "Move 37," a brilliant and unorthodox move played by AlphaGo in its thirty-seventh turn changed the trajectory of the game and helped AlphaGo win. While AlphaGo was a narrow domain-specific, purpose-built model, it was a vivid illustration of the power of reinforcement learning. Lee eventually retired from the game citing AI that "could not be defeated."

There have been tremendous advancements in recent years leveraging reinforcement learning and related methods that allow foundation models to build on strategies that are reminiscent of the human thinking process. These "chain-of-thought" behaviors allow LLMs to try several approaches to accomplish a task, experiment with hypotheses, examine progress and backtrack, etc.

3

Natural Language "Instruction Following"

The final piece that defines the GenAI era is the ability for users to interact with these powerful models in natural language. There is no requirement to learn a programming language or to understand concepts of probability and statistics.

A user can interact with models using the same kind of language they would use while interacting with a coworker.

While this natural language wrapper to interact with a model might seem gimmicky to those who have had the frustrating experience of dealing with "chatbots," it has profound implications. The capability, referred to as "instruction following," allows users to leverage the same general-purpose model for different tasks by providing appropriate instructions in natural language.

By describing what you want to do, one can change the behavior of the model seamlessly and have it perform a task that it had never before been asked to do (called "zero-shot learning").

The description can be augmented with a handful of examples to illustrate what is expected (called "few-shot learning"). Without this ability, general-purpose models would not be nearly as useful.

What's the Catch?

These benefits come with a price—computing resources, and lots of them. As shown in Figure 6, deep-learning approaches (underlying GenAI) keep improving in accuracy with larger datasets and bigger models. In technical parlance, the "model capacity" of deep-learning models to encode information about datasets is far higher than traditional ML models as data size and model sizes increase. And we are yet to hit the limits of such improvements. Foundation model providers are in a race to develop larger and better models. And having already processed the entirety of internet data, newer approaches are relying on existing models to synthetically create better quality datasets by simulating worlds and acting within them.

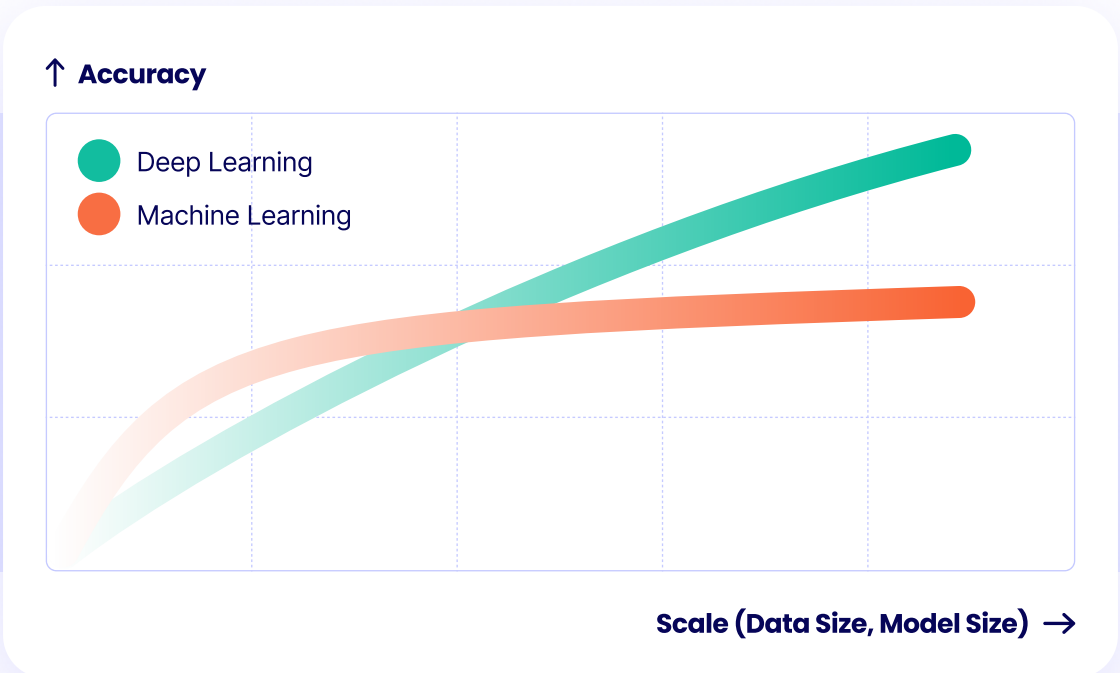


Figure 6: The capacity of deep-learning models is far higher than traditional ML models

But to train, build, and deploy larger models with larger datasets, we need large amounts of computing resources. Researchers in 2009 published work showing how they could leverage graphics processing units (GPUs) commonly used for gaming applications to accelerate computing workloads for deep-learning models.

Since then, there has been tremendous innovation across hardware, infrastructure, and software enabling massive gains in computing efficiency. Regardless of these improvements, the demand for computing resources due to GenAI shows no signs of slowing down any time soon.

Intelligence and Automation

Now that we have a sense for what makes GenAI different from the traditional software or ML automation paradigms, we can see how they compare to one another and to human performance, beginning with intelligence and automation.

While "intelligence" can be defined in a variety of ways, we alluded earlier to the fact that ML systems exhibit intelligence because of their ability to "learn" without being explicitly instructed. A key characteristic of intelligent systems is that they can "adapt" to novel situations not seen before. Automation on the other hand refers to the ability to perform tasks at scale—cheap and fast.

Figure 7: GenAI systems vs. ML systems

Figure 7 summarizes where the various approaches stand in relation to one another. While GenAI systems are more intelligent than ML systems, even approaching close to human performance on certain benchmarks, they are still much harder to scale compared to software or ML systems because of their computing demands.

Autonomy and Predictability: The Path to "Agentic AI"

One of the hallmarks of intelligence is autonomy or agency to act, which is another important dimension to consider. In the context of business, this is the ability to define the "what" and to decide what to do next in a given situation, not just to decide how to accomplish a given task. Software and ML systems are paradigms designed for execution, i.e., the domain of the "how," whereas GenAI allows the possibility of designing systems that can autonomously figure out "what" actions to take toward achieving a larger goal.

However, automation requires reliability. In other words, automated systems work reliably when their behavior is predictable and deterministic. While humans have high agency/autonomy and can also exhibit consistent and predictable behavior, when necessary, autonomy and predictability often are at odds in artificial systems.

Figure 8: GenAI systems are considerably less predictable than ML systems for repeated tasks

Because of their general-purpose nature and because of their ability to accomplish a variety of different tasks without close supervision, GenAI systems tend to exhibit a much higher variation in the nature of their outputs when compared to ML systems and software. In other words, GenAI systems, while capable of successfully accomplishing a variety of tasks, are much less predictable in the exact nature of their responses for any given task done repeatedly.

Can one square this tradeoff to have systems with high autonomy and agency that can also be amenable for automation at scale?

Enter "agentic systems." While we have bemoaned the lack of shared understanding of what "agentic" means, the following is our perspective.

It is now well understood that automation systems—whether they are vanilla software automation, or ML, or smaller GenAI "intelligent" systems designed for narrow tasks—can be engineered to have consistent, predictable behavior by providing appropriate guardrails. What if we could take a large GenAI model, provide it with sufficient context, and give it access to several narrow, purpose-built, automation "agents" that can carry out specific tasks? That is exactly the idea behind an "agentic system"—a powerful, high-context GenAI "orchestrator" that can invoke a number of purpose-built agents for specific tasks. The orchestrator dynamically takes actions as necessary by calling on the right agents to perform tasks. Protocols like Anthropic's Model Context Protocol (MCP) and Google's Agent2Agent (A2A) protocol are attempting to simplify and standardize how such an orchestrator can communicate with agents, push and receive data, and invoke agents to perform specific tasks.

Agentic AI is not about having task-specific agents—those have existed since the dawn of software. It is about the ability to invoke the right ones at the right time and with the right information—it's about what to do and when. In other words, GenAI acts as the decision-maker while using task-specific, purpose-built agents for execution.

It is still very early days, and the field is rapidly evolving. There are many challenges to be solved regarding how a GenAI decision-maker can intake, maintain, and update all the necessary context for accomplishing a high-level goal. One thing is for sure though: the paradigm has already changed, and the future is going to be exciting.

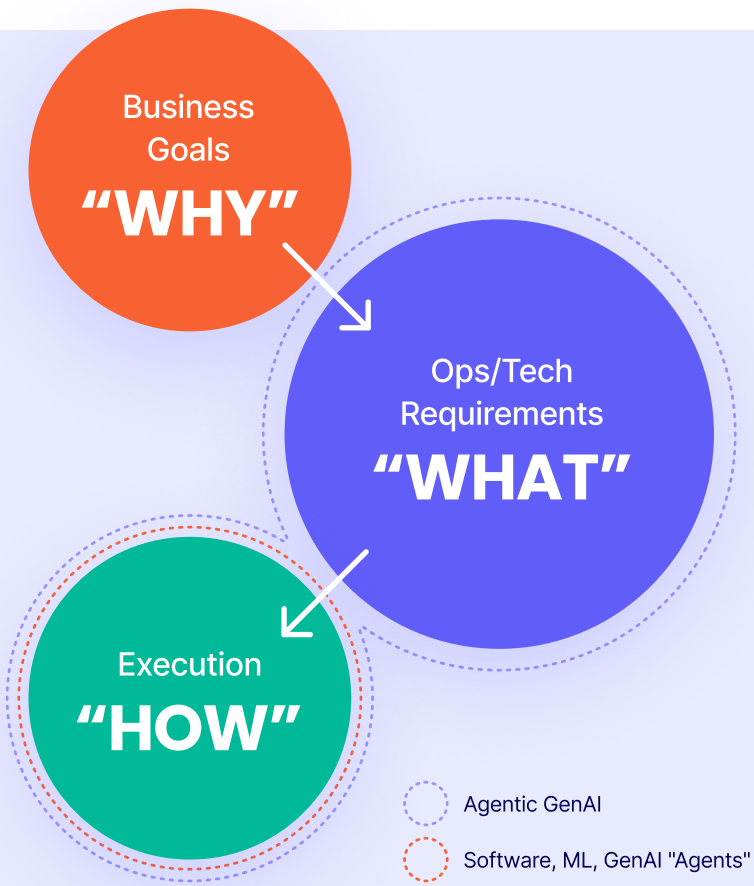


Figure 9: With agentic systems, GenAI is the decision-maker, invoking purpose-built agents to perform specific tasks

Are we close to Artificial General Intelligence, or AGI?

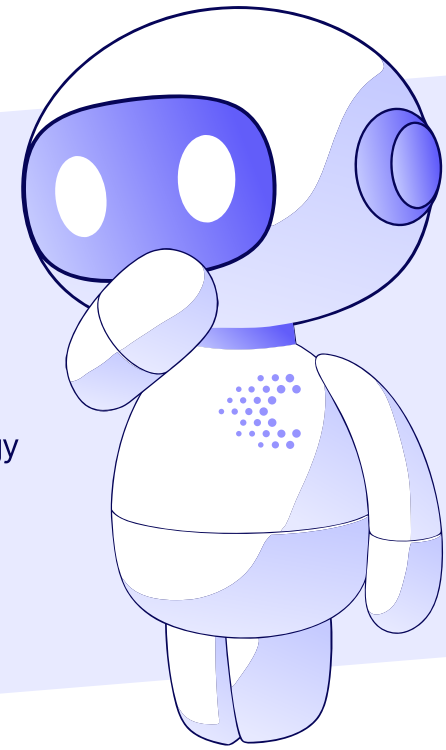
While our concern is with the world of business, automation, and productivity, no discussion of GenAI is complete without addressing whether AI will match or surpass human-level intelligence across the board. This is referred to as artificial general intelligence, or AGI, to distinguish from AI. There have been several claims recently about how AGI is just around the corner and will exhibit "sentience" and "consciousness."

However, the prevalent view among leading AI researchers seems to be that merely scaling up LLMs will not lead to AGI; fundamentally new approaches focused on understanding the physical world, reasoning, and planning will be necessary.

Leveraging models for engineering applications is one thing but attributing sentience or anthropomorphizing models just because they display potentially human-like behavior is a mistake. Renowned philosopher, Daniel Dennett, summarized this almost 40 years ago—far ahead of his time—imagining an LLM-like world with a thought experiment in an essay about the concept of "self":

> *". . . I want to imagine something some of you may think incredible: a novel-writing machine. We can suppose it is a product of artificial intelligence research, a computer that has been designed or programmed to write novels. But it has not been designed to write any particular novel. We can suppose (if it helps) that it has been given a great stock of whatever information it might need, and some partially random and hence unpredictable ways of starting the seed of a story going, and building upon it. Now imagine that the designers are sitting back, wondering what kind of novel their creation is going to write. They turn the thing on and after a while the high speed printer begins to go clickety-clack and out comes the first sentence. "Call me Gilbert," it says. What follows is the apparent autobiography of some fictional Gilbert. Now Gilbert is a fictional, created self but its creator is no self. Of course there were human designers who designed the machine, but they didn't design Gilbert. Gilbert is a product of a design or invention process in which there aren't any selves at all. That is, I am stipulating that this is not a conscious machine, not a "thinker." It is a dumb machine, but it does have the power to write a passable novel."*

GenAI Impact



One of the most common and difficult challenges faced by enterprises in their modernization or digital transformation efforts is the gulf between the "business" and "IT/engineering/technology." Business teams speak an entirely different language than technology teams, and this makes collaborations complex and challenging.

GenAI provides a path for breaking down these barriers.

Trading Off Expertise and Operational Complexity for Computing

Deploying software, ML, or AI automation systems is a significant undertaking for any enterprise. It requires resources with the right skills and expertise, and cross-functional collaboration among other things.

GenAI, because of the factors mentioned above, allows business users to bypass IT and technology teams to experiment with the power of AI. You don't need expensive IT architects, PhD data scientists, or ML engineers to take down requirements, spec a prototype, and develop it for the business teams to see whether a new idea or experiment is feasible. Business users do not have to learn how to program in Python or become proficient with a complex analytics tool; they can interact directly with GenAI using plain natural language. You are trading off this operational complexity and niche expertise around ML/AI for GenAI computing.

This has parallels to the cloud computing transformation where companies moved away from the necessity to have in-house expertise /resources to maintain on-premises IT infrastructure and instead relied on managed cloud services.

Democratizing Intelligent Automation

A direct consequence of this is that the power of AI—previously the domain of data scientists and engineers—is now accessible by business users. GenAI democratizes access to intelligent automation. It can remove bottlenecks and reduce dependency on limited and expensive technology resources. This can happen in a variety of ways, with just a few examples being:

Alleviate skill and knowledge gaps: Users who need to upskill can quickly use GenAI tools or copilots without having to depend on their technical counterparts.

Address bespoke automation needs: Business users can leverage GenAI tools to achieve automation for micro use cases that are too small to involve technology teams but can create a compounding effect in productivity gains.

Provide seamless access to relevant business information and context: Searching for the right information is much easier thanks to GenAI tools such as RAG (retrieval augmented generation) where employees can just ask questions and receive answers instead of having to search across vast internal repositories.

Organize, analyze, and summarize information: GenAI tools can compile information from multiple sources, analyze it, and summarize key takeaways and action items.

Triage and prioritize tasks: GenAI tools can triage and prioritize high-ROI tasks that need human attention for customer-facing teams or operational and support teams.

It is worth noting here that GenAI not only helps business users but can also increase productivity for technology workers. Access to powerful GenAI tools creates an empowered workforce, resulting in efficiencies and productivity gains.

Innovation Velocity

A great thing about an empowered workforce is that it can accelerate innovation across the board. Here again, there are parallels to the cloud computing transformation. Moving technology infrastructure from on-premises to cloud providers allowed teams to swiftly onboard and decommission infrastructure as necessary. This allowed teams to take more calculated risks by being able to quickly vet ideas with uncertain payoffs, which otherwise would have been impossible in the on-premises world.

Similarly, an empowered workforce with access to GenAI tools allows for rapid cycles of experimentation, exploration, and testing of new ideas. The cost of failures goes down, allowing a higher volume of ideas and experiments to be tried and validated. This creates a virtuous cycle of an empowered workforce taking initiative and ownership pursuing continuous improvements.

Is GenAI Going to Replace Technology Workers?

Does this mean GenAI will replace humans across a variety of roles? It seems like every day brings a new sensational headline claiming how many millions of jobs will be replaced by AI.

While automation definitely can lead to the elimination of certain types of roles, we believe fears of AI replacing human workers in the millions are overblown. Agentic AI is still in its very early days and there are a number of technical challenges that need to be solved for it to be effective at scale.

GenAI democratizes access to intelligent automation, but it does not render technology expertise redundant. Software systems, ML systems, and small GenAI models for narrow tasks will continue to be relevant, and building, deploying, and maintaining them will require technical expertise.

And then there are practical, operational, regulatory and other last-mile deployment realities. A good example is the case of automating the work of radiologists. Geoffrey Hinton, Nobel Prize laureate, and often referred to as the godfather of AI, famously suggested almost a decade ago that "we should stop training radiologists now" implying that the job would be automated away soon. And yet, a decade later, radiology as a practice is thriving with demand at an all-time high, albeit with AI-assisted tools.

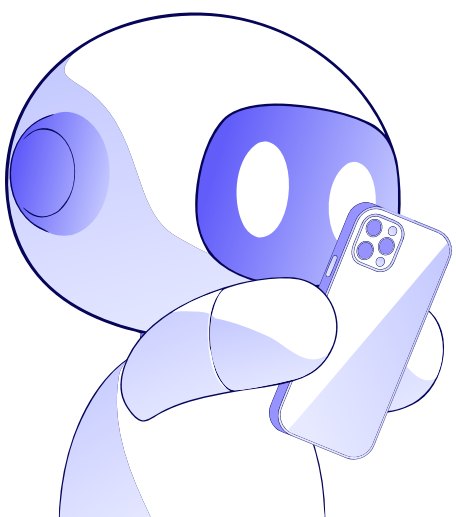
GenAI tools greatly accelerate prototyping and validation, but deploying systems at scale is not something that can be automated away, at least not today with the current state of technology.

Bottom Line: A New Layer of Enterprise Infrastructure

What does this all mean for business leaders? GenAI is the modern-day equivalent of the massive room-size computers of the 1940s when Thomas Watson, the then president of IBM, purportedly famously misjudged the world market for computing to be "maybe five computers." It is hard to imagine any job today that does not rely on a computing device.

The difference with GenAI tools is that they are already seen as necessary tools in an employee's toolbox. In other words, whether you like it or not, GenAI has become a new layer of your enterprise infrastructure.

If employees are not provided with the right GenAI tools, they will "bring their own" access to public tools to help them with their jobs.



When "New" Is Actually Normal

"... has ... become something of a scientific bandwagon.

... has received an extraordinary amount of publicity in the popular as well as the scientific press.

... has perhaps been ballooned to an importance beyond its actual accomplishments."

You may be assuming that the above quotes are about AI today, but they are in fact excerpted from a 1956 article by Claude Shannon on information theory, which was the exciting new technology back then.

And the authors of this essay posit that AI, while transformational, is still a "normal" technology—along with general-purpose technologies such as electricity and the internet—and they contrast this view of AI with utopian and dystopian visions of AI as a superintelligent entity.

New technology is not new. We will get through this era of excitement—with real breakthroughs as well as undeserved hype—but as with transformational technologies, the reality will be very different than what even the best futurists portend.

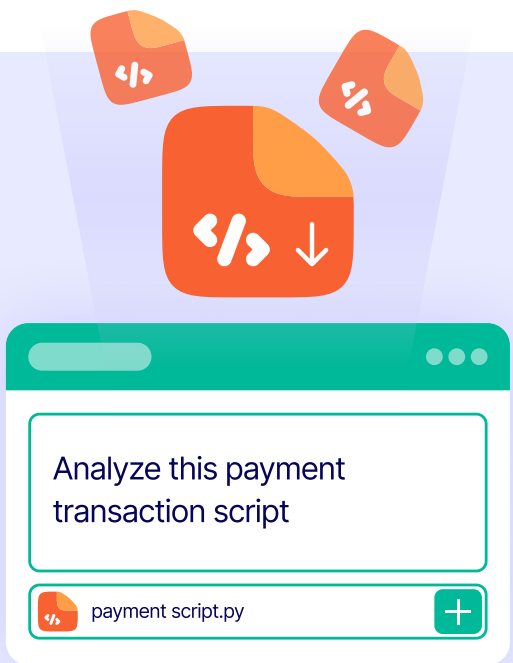
A Familiar Story: The BYOD Parallel

If this feels familiar, it should. A decade ago, enterprises scrambled to respond to the bring your own device (BYOD) trend. Employees began connecting personal smartphones and laptops to corporate networks long before IT could set guardrails. The result was a sudden, uncontrolled expansion of the attack surface. Security leaders had to rush in mobile device management, containerization, and access policies just to regain visibility. That challenge lingers even today, with only partial solutions available to reduce the risk to tolerable levels.

GenAI echoes the risks of BYOD, but the comparison also highlights how much more serious today's challenge is. BYOD risk was tied to devices entering the network or gaining access to data sitting in corporate environments. GenAI risk is about data leaving the enterprise, often invisibly, and at scale. Worse, once sensitive information is provided to a GenAI system, it may be used for training, effectively stripping the enterprise of ownership or control over that data. With a single prompt, an employee could transfer entire datasets, customer conversations, or proprietary code into systems that IT may never see again.

There are also adjacent risks such as the murky question of intellectual property rights over AI-generated ideas or outputs, but those are outside the scope of this discussion.

The BYOD story shows that when convenience outruns governance, enterprises must adapt quickly. The difference this time is that GenAI doesn't just expand the perimeter, it dissolves it. The potential for unintentional leakage, end users unknowingly accepting end user licensing agreements (EULAs), obscured fine print in vendor agreements, and contamination of training pipelines makes today's risks exponentially greater than those of a decade ago. And unlike BYOD, which primarily hit regulated industries first, GenAI impacts every organization, across every sector, and all at once.



Collateral Repercussions

While we have reviewed the transformative power of GenAI, we know from the philosopher Voltaire (Or Uncle Ben in the Spiderman movie) that "with great power comes great responsibility." Imbibing GenAI tools with "responsibility" is a challenge and remains an active area of research and development. And we are only just beginning to understand emerging risks around deliberate scheming from some of the frontier models.

This paper highlights some of the potential unintended consequences of GenAI as they relate to data.

Software and ML automation paradigms, by design, need carefully constructed and curated inputs and often take the form of structured datasets. The inputs, outputs, and any other information collected are specified upfront, and this makes the problem of data governance and data security contained and tractable. For example, Figure 10, which has been adapted from a widely cited paper from Google published in 2015, shows components of a typical ML deployment.

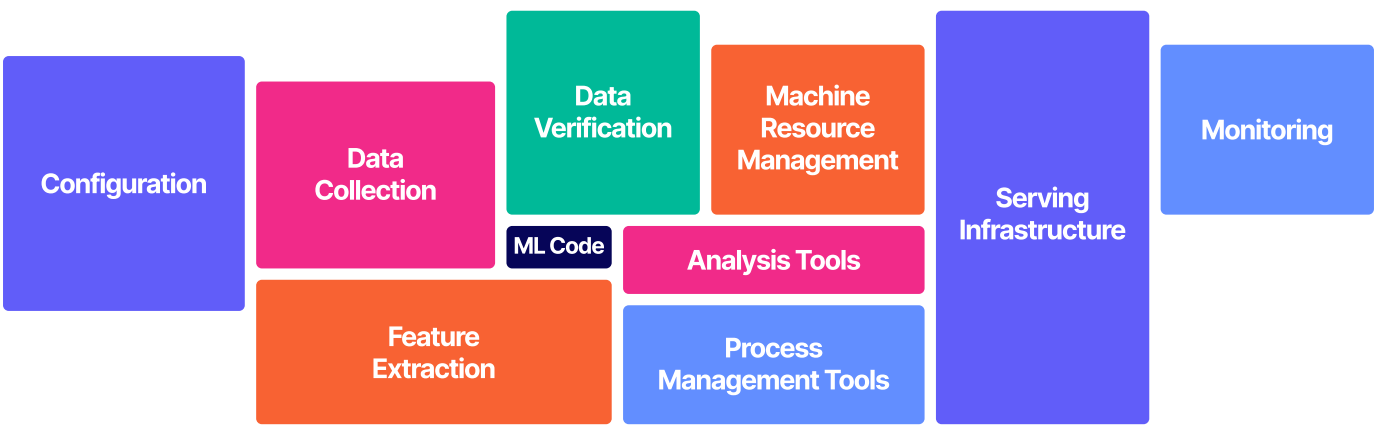


Figure 10: Components of a typical ML deployment

Because of all the upfront effort, one can track the nature of data that is consumed, processed, and created by these systems; identify whether any of it is sensitive; and determine how to secure such data and their flows appropriately for authorized users.

Upending the Principle of Least Privilege

Least privilege access is a security model where any entity, whether it be user, program, or process, is given only the minimum level of access and permissions to resources necessary to perform its legitimate functions—and nothing more. When you have traditional software automating defined tasks, this is often accomplished by providing users and employees role-based access controls (RBAC). Although, in practice, most organizations struggle to implement them effectively.

GenAI however exacerbates the struggle by upending the entire paradigm: least privilege itself becomes a constraint that conflicts with the very way these systems are designed to operate.

GenAI works better when it has access to more data, and enterprise GenAI tools tend to perform better and lead to higher productivity gains when they have access to more business data and business context.

Because of the very nature of GenAI, which allows users to leverage it in a variety of ways, users continue to find new applications of GenAI, most of which emerge from organic experimentation and curiosity, rather than top-down, business-driven planning.

If an entity cannot be defined by the nature of tasks it will be used for or the types of data it needs access to, it becomes infeasible to set up least-privilege access permissions. In addition, a user may have appropriate access to a dataset and legitimately provide it as input to a GenAI tool.

But once that data is ingested, it is no longer bound by the user's original permissions. Instead, it can be absorbed into the model, surfaced in future outputs, or even become accessible to others leveraging the same tool. In short, GenAI does not necessarily inherit the access controls of the data, rendering least privilege effectively unenforceable.

The Expanding Risk Surface

GenAI creates a vast and ever-expanding data surface, complicating enterprise data governance and security in several interconnected ways:

- **Input leakage**
- **Output exposure**
- **Accessibility without oversight**
- **Second-tier supply chain risk**
- **Governance gaps in training data**
- **Application code risk**

Input leakage

GenAI can ingest data in its raw form, across various types including text, images, audio, video, structured data, and more. End users can now direct GenAI tools to new datasets with minimal effort or expertise. Instead of being limited to carefully curated, structured tables with defined schemas and relationships, these datasets may include sales call recordings, CRM email notes, customer service transcripts, and more. In practice, employees are feeding prompts with highly sensitive business information, from customer PII and intellectual property to financial forecasts and even source code.

Many tools retain those inputs; some reuse them to train future models; and the fine print that governs those practices is often buried in dense EULAs that most people never read.

Output exposure

Generative models don't just consume, they synthesize. A prompt can unintentionally pull insights from across datasets and expose them to users without proper clearance. In some cases, outputs can even "hallucinate" data that appears legitimate but embeds fragments of real, sensitive training material.

Gen AI tools work better when they have better context about what they are being asked to do. Not only is Gen AI ingesting existing data, but users are creating new data in the form of extensive, detailed prompts, documenting business context, internal processes, and other potentially sensitive or business-critical information to guide GenAI.

Accessibility without oversight

Where traditional enterprise systems required vendor onboarding and IT provisioning, GenAI is everywhere: embedded in office suites, browsers, chat tools, and SaaS platforms. Employees can adopt it instantly, bypassing governance entirely. This frictionless access is the engine of "shadow AI," and every unsanctioned use of GenAI is a potential data exfiltration event happening invisibly, at scale, and outside your governance perimeter.

Governance gaps in training data

Once data enters an AI model, control effectively ends. Enterprises cannot easily retract or govern how their information is used. Proprietary knowledge may persist in weights or embeddings, surfacing in future outputs long after its source has been forgotten. Have you ever seen a form in any GenAI tool that allows you to submit a request to remove certain information that has been ingested—similar to what you see in privacy regulations such as the General Data Protection Regulation (GDPR) or the California Consumer Privacy Act (CCPA)? We have yet to see this and doubt it's coming until regulation drives the change.

Second-tier supply chain risk

Even when a vendor looks secure, they may rely on their own subcontractors like cloud hosts, annotation services, or third-party AI labs. Each introduces its own EULA and policies. Sensitive enterprise data can ripple through multiple unseen hands, but accountability remains squarely on the enterprise. You could have an existing vendor that previously completed your onboarding process but now uses a GenAI tool that will allow your data to be the training data, and you are not aware of that downstream impact.

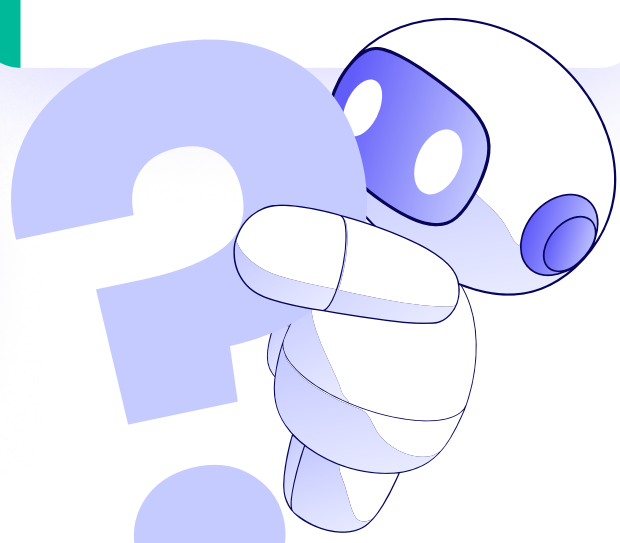
Application code risk

AI is increasingly writing the code that underpins business systems. Developers using GenAI like Microsoft Copilot to generate code may unknowingly import insecure dependencies, propagate vulnerabilities, or embed code under conflicting open-source licenses. Once deployed, these weaknesses become embedded in the software supply chain itself.

Together, these dynamics make GenAI fundamentally different from past technology waves.

GenAI raises challenging questions:

- Who is at the other end of this new infrastructure layer, and what can they see?
- What happens when users ask questions about information that they are not supposed to be privy to?
- What sensitive information might be going into these GenAI tools?
- What if users use GenAI to try and bypass access restrictions to get such information?
- Is any sensitive information retained by GenAI tools that can be inadvertently "remembered" later?
- What forensic evidence is available for GenAI use in case of a breach or misuse?
- Who else within the enterprise has access to GenAI interactions and sessions?
- What assurance exists for the use of GenAI with vendors and third parties?
- What happens to GenAI outputs? How are they being shared within and potentially outside the enterprise?
- How does GenAI use align to regulatory and contractual obligations?
- How do you govern deviations from expected outputs?



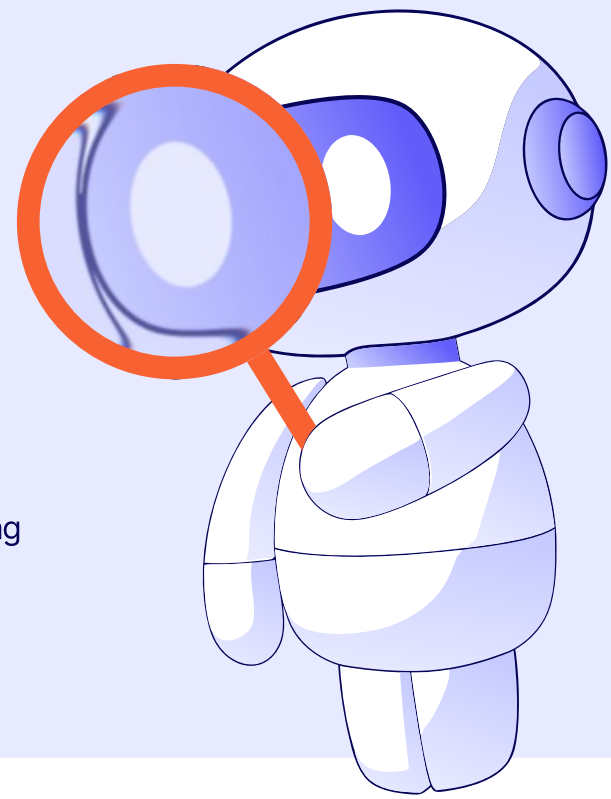
Data Security Risk Governance

The GenAI era has expanded data risks in ways that traditional frameworks, access controls, DLP, and compliance checklists do not fully address. Enterprises need to focus on building out a mature data security risk governance program that treats data as an asset and measures exposure in quantifiable, business-relevant terms.

This requires more than cataloging threats. It means adopting risk quantification techniques that can express data risk in measurable units, which boards and regulators can act on. For example, financial risk allows leaders to weigh trade-offs, compare scenarios, and make faster, smarter decisions.

Effective quantification depends on new inputs beyond traditional metrics:

- Data sensitivity and context (IP, PII, financials, source code)
- User behavior and intent (benign use vs. shadow AI misuse)
- Regulatory and contractual obligations
- Data element metadata, such as unique count of personal data and toxic clustering
- Access to data using access controls or exploitation of vulnerabilities



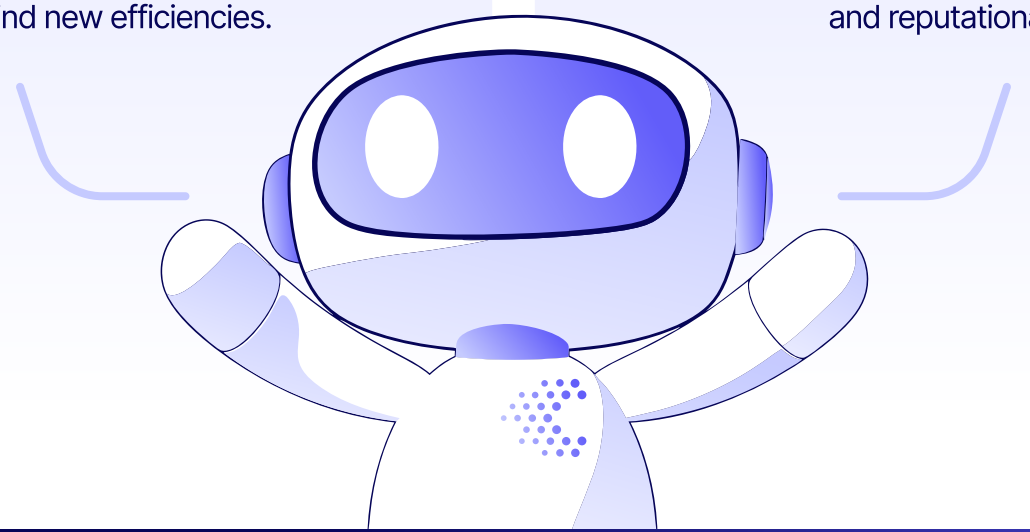
The specifics vary by industry and appetite for risk, but the principle holds: measure what matters so boards and CISOs can allocate resources intelligently. Done well, governance becomes less about slowing innovation and more about ensuring it can happen safely and sustainably.

Why the Cost of Inaction Is Higher

If enterprises hesitate, they face a dual bind.

On the one hand, ignoring GenAI means losing competitive ground as empowered employees and business teams find new efficiencies.

On the other hand, adopting without governance means risking multimillion-dollar breaches, regulatory penalties, and reputational harm.



IBM's numbers tell the story: organizations that had fully deployed AI-driven security automation saw breach costs reduced by an average of **$1.76 million** compared to those that had not. In other words, AI is both the risk vector and the most effective shield.

Executives cannot afford to view GenAI as a passing experiment. It is already embedded into enterprise workflows, often invisibly. As with BYOD, adoption is inevitable. But this time, the stakes are far higher.

The Path Forward: Context-Driven Visibility and Layered Controls

The question, then, is not whether to adopt GenAI, but how to do so responsibly.

Every new technology comes with new vulnerabilities, threat vectors, and risks. We have seen that GenAI is a de facto layer of the enterprise infrastructure and thus the foundational principles of identifying, monitoring, and mitigating risks apply. The answer lies in context-driven visibility, detective monitoring, and preventative controls. A layered defense that mirrors the BYOD playbook is required, but it has to scale for today's data-driven risk. Last, legal and regulatory frameworks must move faster at driving controls around sensitive data.

There are two ways to look at it – an inside-out perspective from the enterprise point of view and an outside-in perspective from the infrastructure point of view.

Inside-Out Governance (Data Hygiene)

The foundation is visibility into what data exists, where it resides, and who can access it. Enterprises must categorize and classify sensitive assets including employee and customer PII, proprietary IP, and confidential code, and they must remediate overly broad permissions.

Data loss prevention (DLP) starts at what the data is, and who has access, long before DLP controls are implemented for data in motion. Shadow data should be purged, duplicates eliminated, and retention policies enforced. Without this hygiene, every GenAI interaction is a potential leak.

Data categorization, which is the process of organizing related data into categories and even sub-categories, should be used over classification as it is more powerful and can create flexible controls for empowering the business rather than hindering it. This is because it allows departments or people the use of certain data types as opposed to a large set of data based on a classification; for example, the legal team can use a specific GenAI tool, with legal contracts.

Outside-In Controls (GenAI Monitoring)

While good data security governance practices can address and mitigate data risks in the long term, we need to augment that with appropriate controls on the GenAI infrastructure layer.

Enterprises require visibility into GenAI use itself. They need to see which tools are being accessed, what prompts are being entered, and whether sensitive data is leaving. That visibility can be enabled through browser extensions, proxies, or API-level integrations.

From there, organizations can apply detective and preventative controls. Detective controls monitor prompts and outputs in real time, flagging risky sessions or anomalous data flows. Behavior analytics and anomaly detection, especially in the agentic world, combined with data context will be key in detecting and identifying such risk. Preventative controls block unsanctioned tools, filter sensitive prompts before they leave, de-identify sensitive data as it is entered into prompts, and enforce role-based restrictions on AI-driven insights.

Together, this creates a feedback loop of trust: employees can use GenAI safely, security teams retain oversight, and business leaders gain confidence that innovation won't come at the cost of exposure.

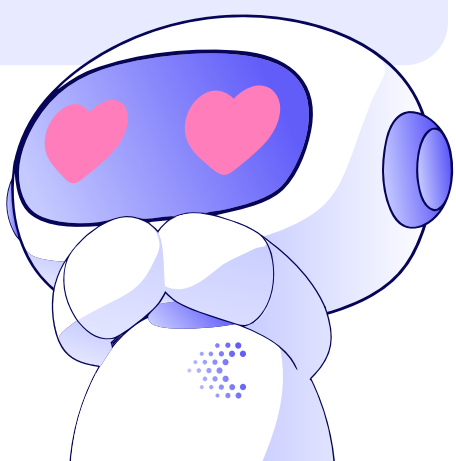
Identifying What Is Under Your Control

To illustrate the inside-out perspective, we'll share this account of a security team at a dating service that needed to detect and combat malicious, spammy, and undesirable behavior on their platform.

Initially, they used an outside-in approach, but this was challenging because the problem space was unbounded. Whenever they succeeded, external actors adapted with new behaviors, resulting in a never-ending game of catch-up.

So, they changed tactics. Instead of focusing only on all the ways malicious actors could act, they focused on what was within their control, which was the ability to define what was benign, safe and "desirable" behavior on the platform.

They built algorithms to identify and promote such behaviors and finally started making a dent in combating the unwanted actors.



The Regulatory Horizon

The governance challenge is not only internal. Regulatory and legal expectations are accelerating, and enterprises will soon be required to prove how they manage the risks of GenAI. The EU Artificial Intelligence Act (EU AI Act) is the clearest signal: it explicitly classifies certain applications of AI as "high-risk," attaching strict requirements for transparency, accountability, and human oversight. Penalties for noncompliance are designed to be punitive, measured in percentages of global revenue rather than flat fines. For multinational companies, that transforms AI governance from a best practice into a legal necessity.

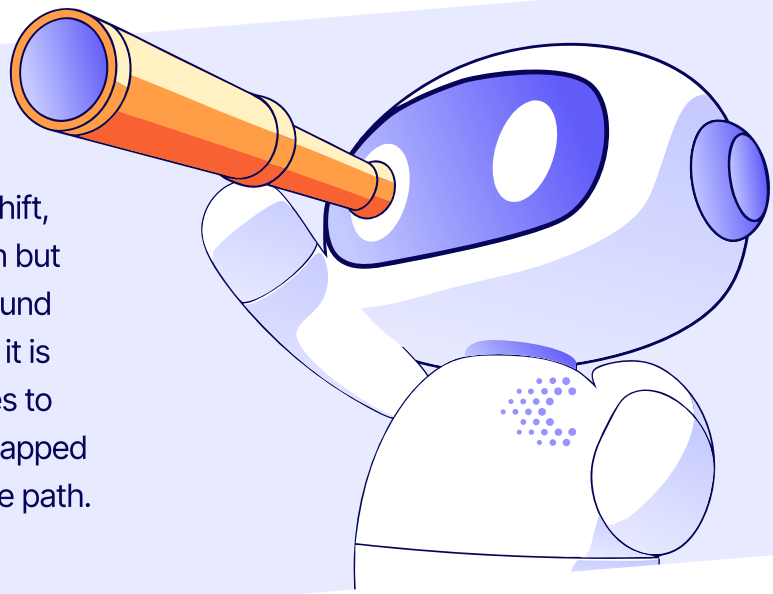
In the United States, the Federal Trade Commission (FTC) has made it clear that AI use will be scrutinized under its mandate to prevent "unfair or deceptive practices." That means enterprises cannot hide behind technical complexity: if customers or employees are misled about how their data is handled, or if AI introduces undisclosed risks, executives may be held accountable. Early enforcement actions have already shown that regulators are willing to make examples of companies that move too fast without controls.

At the state level, California and New York are leading a wave of AI and privacy legislation that builds on existing frameworks like the CCPA. These laws converge with broader privacy obligations, meaning that enterprises must treat AI governance and data privacy as one unified program.

The collision is already visible: today, a consumer can request under GDPR or CCPA that their personal information be deleted from a company's systems. But there is no equivalent mechanism to request removal of data once it has been ingested into a GenAI model. The absence of a "delete-my-data-from-your-model" process will soon become a flashpoint between regulators and enterprises.

For executives, the lesson is clear: regulatory pressure is moving faster than many expect, and waiting for standards to stabilize is not a viable strategy. Governance frameworks must be implemented now, not just to reduce risk, but to demonstrate to regulators, customers, and boards that AI adoption is happening intentionally and responsibly.

Looking Ahead



GenAI represents a generational paradigm shift, one that democratizes intelligent automation but also dissolves long-standing boundaries around enterprise data. If BYOD taught us anything, it is that blocking adoption only drives employees to use unsanctioned solutions. Enablement, wrapped in visibility and control, is the only sustainable path.

But unlike BYOD, the data security risk with GenAI is far larger. A misplaced device could expose files; a misplaced prompt can expose the crown jewels of an enterprise's intellectual property. The consequences are amplified, the attack surface broader, and the governance stakes higher.

The enterprises that succeed will be those that recognize that this is not about approving one more system, it is about governing an entirely new layer of enterprise risk and opportunity. Governing it requires context-driven visibility, detective and preventative controls, and a mindset that security is not a brake on innovation but the foundation that makes innovation safe and enables the business. The companies that master AI governance will innovate faster, attract more customers, and win regulator trust. The others will spend years in pilot purgatory, burning capital while their competitors compound advantages.

Every innovation is a potential vulnerability, and every vulnerability a potential headline. Enterprises that learn this lesson the fastest will be the ones still leading when the hype cycle clears. GenAI dissolves the perimeter. The leaders who understand this will define the future of secure innovation.