databricks

The New Architecture Of Agentic AI

How CIOs, CTOs and CDOs Are Rebuilding the Foundation for the Agentic Era



A Databricks Executive Lens Publication | 2026

Contents

Chapter 1: The Inflection Point 03

Chapter 2: Stop Asking What AI Can Do. Start Asking What AI Should Own 08

Chapter 3: Your AI Is Only as Smart as Your Data Layer 12

Chapter 4: The Gap No One Warned You About 15

Chapter 5: Governance Is Not the Brake. It's the Engine 19

Chapter 6: The Operating Model Is the Strategy 24

Chapter 7: Five Strategic Plays for the Next 18 Months 29

What You Build Now Determines What You Can Do Next 32

About Databricks 33

Contributors 34



The executives in these pages are not describing a future that might happen. They are describing a transformation that is happening now, in organizations they lead and organizations they serve. The gap between the companies building this new architecture and those still debating whether to build it is widening faster than most organizations realize.

What separates them is not budget, talent or access to better models. It's a series of connected decisions about data governance, semantic layers, intelligence infrastructure, agent lifecycle management and workforce redesign. Together, these add up to something structurally different from what organizations have built in the past. And that foundation determines the options available to organizations for agentic AI. This eBook documents what that foundation looks like, who is building it and what they know that others do not yet.

Chapter 1

The Inflection Point

Most enterprises do not fail at AI because they lack talent or ambition. They fail because yesterday's data and application architectures were never designed for systems that reason, decide and act. After years of experimentation, the organizations pulling ahead have stopped asking what AI can do. They ask what decisions and workflows AI should own. And then they build the infrastructure to make that happen.

> "There is more AI activity than AI value right now. The difference between companies that generate meaningful outcomes and everyone else comes down to one thing: they start with the outcome they want to drive and work backward. They are not chasing the technology."
>
> — Arsalan Tavakoli-Shiraji, Co-founder and SVP, Field Engineering, Databricks

Naveen Zutshi, CIO of Databricks, has watched this shift accelerate across every sector. He's seen healthcare systems reducing administrative burden for clinicians, financial institutions generating alpha through more sophisticated techniques and manufacturers transforming operations end to end. And yet the gap between experimentation and creating value has never felt wider. Something is broken in the path from AI idea to AI impact, and it's not the models — it's the architecture, operating model and governance framework.

The diagnosis underneath all three is the same: the models are smart enough. What they do not have is context. A frontier model has no native sense of what your company means by "active customer," which version of a margin number was flagged because of a one-time promotion or which policy revision applies to which region. Without that context, even the most capable model returns confident, generic answers to specific business questions. With context, the same model can reason against your business reality and produce answers that hold up under scrutiny. The strategic question is no longer which model to choose but how to give the model an accurate, governed view of the enterprise it's operating inside.

43%

Many firms are using or actively rolling out AI across departments — but only 43% formally require teams to measure the business impact.

Economist Enterprise, "Making AI Deliver," 2026

The Economist Enterprise research is sobering: most firms take between 7 and 12 months simply to move a project into live production. Which means for an AI initiative to justify its investment, returns must materialize from the moment something goes live. That pressure is forcing a different kind of cadence. And it reveals that the real cost burden of running AI at scale falls not on models but on moving and maintaining data, and on the staff hours required to verify what AI systems create.

Which is why control has become the new ceiling on speed. The yardstick that matters now is not how many agents an organization has deployed but whether it can govern, audit and retire them. The leaders pulling ahead govern AI across its lifecycle and not at the approval gate. They monitor model behavior before, during and after a launch. They give agents defined authorities, observability and a clear path to shut something down. And they treat those activities as platform requirements rather than nice-to-haves. Speed can be sustained because of how control is engineered into the foundation.

What is shifting fastest is the question companies are even asking. Kristy Mayer-Mejia, Global Head of Data Transformation at Kraken — the AI-powered operating system used by some of the world's largest utilities to manage more than 90 million customer accounts across 27 countries — describes a noticeable change in client conversations. The work to get data into a usable state is still real and still takes time. But the question is shifting from "How do we get our data ready?" to "Now that the foundation is there, what becomes possible?" And the answer keeps expanding.

> "AI has been a forcing function. Clean data, documentation and context have always been needed. AI just made it nonnegotiable. Clients who would have considered a monthly report a success two years ago are now running hourly dashboards with predictive models on top — and looking quickly toward broad agentic AI. The ones who invested early in their data capabilities are the ones moving fastest now, and the gap between them and everyone else is only going to widen."
>
> — Kristy Mayer-Mejia, Global Head of Data Transformation, Kraken

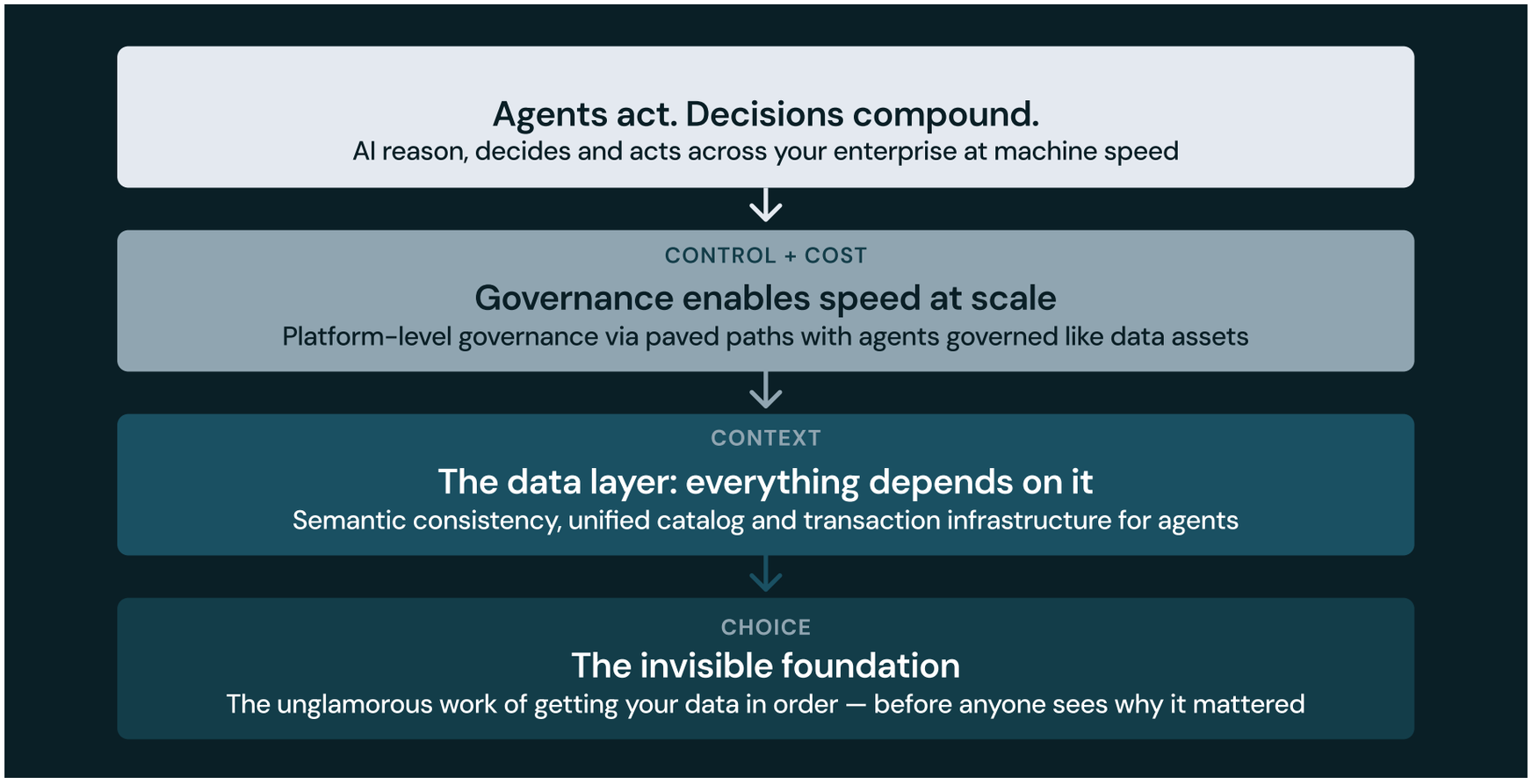
Three questions that separate AI leaders from laggards

1. Have you moved from “What can AI do?” to “What decisions and workflows should AI own?”
2. Can your AI reason against your business reality — your definitions, your governed data, your current policies? Or is it returning generic answers to specific questions?
3. Can you govern, audit and retire the agents you have deployed? Or is your control framework still designed for data and not agents?

Where AI value compounds — and what holds it up

The architecture of enterprise AI is a stack of dependencies. What gets demoed sits at the top — agents reasoning, deciding and acting at machine speed. What makes that visible value possible, and what makes its returns compound, sits in the three layers underneath.

Those three layers are how an enterprise answers the four demands agentic AI puts on every architecture decision: the context for agents to reason well, the control — and cost discipline — to let them act safely at scale and the choice to stay free of lock-in. Get the stack right and all four follow.



Chapter 2

Stop Asking What AI Can Do. Start Asking What AI Should Own.

The shift happening at the leading edge is not about better models but about a fundamentally different question. Most organizations started their AI journey by finding a model, picking a use case and running a proof of concept. That produced demonstrations. This satisfied the need for a demo, but not operational capabilities. The organizations closing that gap have moved from capability-first thinking to outcome-first thinking. And engineered backward from the outcomes to the architecture.

When Barracuda rebuilt its managed XDR detection product, the team did not start with a model. They started with the outcome they were trying to deliver and worked backward, starting with the data layer. They normalized the schema across email, network, application and managed XDR telemetry so that machine learning models and agents could operate with full context. The pattern now extends to web application security, automated configuration, API security and bot protection. The strategic decision Neal Bradbury, the company's Chief Product Officer, identifies as the inflection point: AI was built into the architecture of the product, not added on top of it. That investment in the foundation compounds across the portfolio rather than getting rebuilt with every new feature.

> "Whether intelligence is built into the core of what you ship, or layered on top — that determines how fast it can evolve when the landscape shifts again."
>
> — Neal Bradbury, Chief Product Officer, Barracuda

The conversation about model quality is essentially over. Frontier models have commoditized, and what matters now is what organizations build on top of them — and how fast. The customers doing this properly are going six times faster into production than the rest, according to Kunal Taneja, AVP for Field Engineering in APJ at Databricks. He sees that the curve steepens for the organizations still treating model selection as the strategic question.

> "The agentic evolution is the same as the dotcom revolution. Every industry had to adopt the internet. The same thing is going to happen here — even regulated industries will eventually come down to unit economics. When your competitor operates at a fraction of your cost, regulation stops being the binding constraint."
>
> — Kunal Taneja, AVP for Field Engineering, APJ, Databricks

Outcome-first thinking comes with a second discipline most organizations skip: picking the right workflow to start with. Not every process is ready for an agent. If a workflow has messy inputs, unclear rules and no clear definition of a right answer, putting an agent on it does not make it better. It just makes the wrong answers come faster — and makes it harder to figure out why.

The workflows worth going after first look the opposite way — the inputs are clean, the rules are written down, there is an objective definition of what good looks like and getting it wrong has an actual cost the business can measure. Pick those workflows first and the foundation work starts to pay off. Skip that step and the foundation work becomes a tax with nothing to show for it.

Financial crime compliance fits the profile quite well. The workflows where agentic AI delivers earliest are the ones that are simultaneously the most cognitively intensive and the most precisely documented. Financial crime compliance — know-your-customer processes, anti-money-laundering workflows — requires synthesizing information from dozens of sources against extensively documented rules.

That, says Junta Nakai, VP of Financial Services at Databricks, is the definition of an agent-ready workflow. The financial crime analyst becomes exponentially more effective because the agent does the synthesis. The analyst keeps the ability to judge those outputs.

The work that enables agentic AI — clean data, strong identity and access management, well-documented APIs — is unglamorous by nature. It is also where competitive advantage quietly accumulates. Craig Wiley, VP of AI at Databricks, is direct with his prerequisite: getting a high-quality model is the easiest part of the problem. The hard part is knowing where your data lives, how to connect it and how to build governance that understands not just your data but your agents — what permissions they hold, which systems they are reaching into, how many of them are running in parallel. The organizations willing to do that work before they need it are the ones with the most options when it matters further in the development cycle.

None of this works without evaluations built in from the start. Maria Zervou, Chief AI Officer, EMEA, at Databricks, shares that you do not ship an application unless you know how to evaluate it, and not just from a technical perspective. Evaluation must happen from a business perspective, with business stakeholders involved, before anything goes into production. The most important questions are semantic ones. Does the output mean what the business needs it to mean? Does the system understand what a specific term means in this company, at this stage of a customer relationship? If the answer to either of those questions is no, the model is not the problem.

> "If you cannot trace what data the AI used to answer, you are not secure. And if you are not secure, you will risk regulatory penalties."
>
> Maria Zervou, Chief AI Officer, EMEA, Databricks

The fastest path to a measurable win

Start with a specific outcome: a workflow you want to own, a decision you want to accelerate. Don't look for a model or start with a cool use case looking for a problem. Pick the workflows that simultaneously require a high cognitive load and that are bound by rule-based operational workflows at the same time. Think about a workflow where an agent can do the synthesis while a human keeps the judgment. Then build in evaluation from day one, with business stakeholders defining what a correct answer looks like before anything ships. And treat the unglamorous foundation work — clean data, governed access, well-documented APIs — as the prerequisite for everything that comes next, not the cleanup after.

Chapter 3

Your AI Is Only as Smart as Your Data Layer

AI cannot deliver trusted outcomes when every team works from different definitions of the same metric. Raela Wang, Director of Field Engineering at Databricks, identifies three data-layer failure modes that she sees repeatedly blocking organizations from getting AI into production.

The first is semantic fragmentation: the same metric defined differently across teams. Revenue means one thing in finance, something slightly different in sales ops and something else in the product dashboard. This fragmentation existed long before AI, but AI amplifies it: now you are generating answers at scale from conflicting definitions, and nobody is checking the math.

The second failure mode is data sprawl: the same data living in multiple places, each copy updated on its own schedule by different teams with slightly different logic. When someone asks for the customer count today, what should be a simple answer becomes a question: which version of the data are you looking at?

The third is governance that was not built for AI. Most organizations implemented data governance for the reporting era. That architecture does not scale to AI.

Nick Eayrs, VP of Field Engineering, APJ, at Databricks, describes the transition bluntly: legacy business intelligence (BI) is a world of static dashboards and predefined reports. When you want to explore something deeper, you need specialist support. AI changes everything because it is not static and not just descriptive. You can ask why something happened, what you can do about it and reason across far more data than any analyst ever could. The semantic layer — the governed set of business definitions, certified metrics and data lineage that sits beneath AI — is not a technical nice-to-have. It's the foundation on which every AI output rests.

The cost of skipping that foundation work shows up later, and it shows up everywhere. Models trained on inconsistent definitions produce inconsistent answers. Agents reaching into sprawled data return whichever version of the truth they happened to find first. Governance frameworks built for static dashboards cannot tell you which model used which data to produce which output, and that is the question regulators and auditors are about to start asking. None of these are model problems — they are data layer problems, and the organizations treating them as model problems are about to spend a lot of money learning the difference. The strongest foundations are also the most open. When the data layer is built on open formats, it stays portable across clouds, engines and models — so no single vendor's roadmap becomes your ceiling, and the architecture preserves choice instead of closing it off.

> "Don't build AI on a broken foundation. The data layer is the strategy. The winner is going to be whoever has the strongest foundation. Do the legwork."
>
> — Stephen Ecker, Chief Data Officer, Trinity Industries

Stephen Ecker, Chief Data Officer at Trinity Industries, has lived this transition firsthand. Trinity, one of North America's largest railcar manufacturers and lessors, ran a data environment in which workloads bounced between Azure, AWS and on-premises. The company had nearly 600 distinct measures across the business — and many of these measures were calculated from the same source but with different filters baked in. It was logging 11,000 hours a month inside increasingly bloated dashboards.

The migration to a unified lakehouse architecture took close to a year, with another six to eight months of cleanup. The payoff: an ETA prediction model 50% more accurate than the industry's own benchmarks, agents now interacting with over $1 billion of supply chain procurement and a 15% lift in on-time material delivery.

A generic AI assistant is trained on public data — it has no idea what your business means when it says "revenue," or "churn" or "active customer." The difference, as Ari Kaplan, Global Head of Evangelism at Databricks, puts it, is that a system grounded in your company's specific terminology, business rules and governed data does not just sound confident. It gives you answers you can actually act on.

Three data-layer questions to answer before your next AI investment:

1. **Semantic consistency.** Ask three departments for the same number. If you get three different answers, your AI will give you a fourth.
2. **Single source of truth.** Ask how many copies of your customer table exist, and on what update schedules. The number is almost always higher than leadership thinks.
3. **AI-era governance.** Ask whether your current framework can tell you which model used which data to produce which output. If it can't, you have governance built for reports, not for AI.

Chapter 4

The Gap No One Warned You About

There is a failure point that organizations only discover when they try to take an AI application from prototype to production. The model works. The agent logic is sound. The governance is in place. And then — at scale, under real load, with real users — it breaks. The problem is not the AI. The problem is the database underneath it.

Traditional transactional databases were designed for a world where they supported large, monolithic applications that were designed, built and maintained by humans for predictable and bounded workloads. This one big system managed the application state, operational data and transactions. That architecture made sense when humans were the ones building and running software, says David Hackett, VP of Field Engineering at Databricks, who has spent more than 25 years in data platform infrastructure. It does not survive the shift to agents.

Agents change the equation entirely. They do not operate like human-scale applications. They spawn parallel instances, explore multiple paths simultaneously and move far faster than any human workflow. The atomic unit of a transactional store is no longer a single monolithic database. It is often one store per agent, holding only the context that specific agent needs. And the number of transactional stores being created by agents will dwarf those created by humans. Most legacy transactional platforms were not built cloud-first. They were ported to the cloud, and the architectural decisions made for on-premises environments are still visible — and still creating friction.

Most organizations get tripped up by what this means in practice. Agents need to do what developers do with code: experiment with changes and ideas safely. The right framing, says Shanku Niyogi, VP of Product Management at Databricks, is to treat production data the same way engineering teams treat production code — something you can clone, branch, checkpoint and revert without breaking what is running. An agent exploring options needs to branch the data state, try something, evaluate the result and roll back if it is wrong. Without this freedom to experiment safely, the promise of agentic AI collapses into a more expensive version of what you already had.

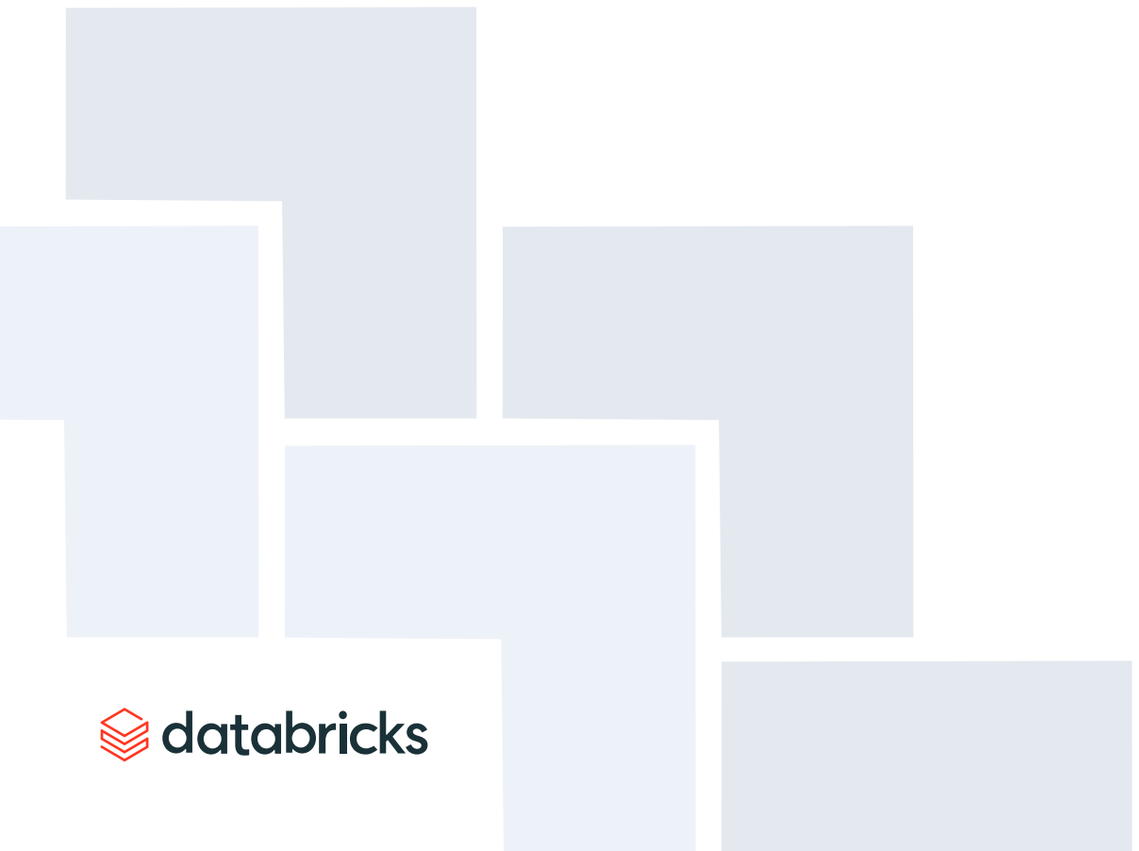
The same architectural intuition plays out in domains where AI must operate against current data, not yesterday's. Nader Mherabi, Chief Digital and Information Officer at NYU Langone Health, runs real-time clinical decision-support models that flag conditions that may have been missed before a patient is discharged from the emergency room. Retrospective data is useless in this scenario where real-time insights are needed. The architecture has to deliver current data to the model fast enough for the recommendation to land in time to matter. Nader's discipline goes further upstream than most: NYU Langone fixes data quality at the transactional systems themselves — the electronic health record and the ERP — so the intelligence layer is not constantly filtering out problems introduced earlier in the pipeline.

> "If you want clean water in your intelligence layer, you have to fix it in the transactional systems first. Filtering dirty water is expensive. Some things you'll still have to filter along the way — but the principle should be to get it right upstream."
>
> — Nader Mherabi, Chief Digital and Information Officer, NYU Langone Health

The most common mistake in the field is defaulting to the existing transactional database before defining what the application actually needs to do. The existing database is familiar, already licensed, already known to the team. For the first agent application, it may be enough. But as the number of agents grows, the gaps compound — scaling, replication, branching, cost of operations — until the architecture that felt safe becomes the bottleneck.

The discipline that avoids this is the work no one actually sees: start with the actual requirements of the agent workload, then map those requirements to what you currently have and look honestly at the deltas. The deltas are where you discover whether your existing platform is adequate or whether the workload needs something purpose-built for AI-native applications.

There is a second consequence that compounds the first. When transactional systems sit apart from analytical and AI workloads, the integration overhead is permanent. Every time an agent application succeeds — and you want to understand why, and replicate it — you face the same stitching problem again. The organizations that avoid this build on a unified foundation where transactional activity, analytical workloads and AI operate on the same platform, under the same governance. When an agent succeeds, the data is already where you need it to run the analysis. No custom pipeline. No duplication. No governance gap.



Four questions to ask before your next agent goes into production

1. **Was your transactional database born in the cloud, or ported there?** If it was ported, the architectural decisions made for an on-premises world are still embedded in how it handles scale, replication and concurrency. Agents find those limits faster than humans do.

2. **Can an agent branch the data, try something, evaluate the result and roll back cleanly?** If not, your agents are not running experiments — they're running bets. And that is not a safe operating model at scale.

3. **Are you fixing data quality at the source, or filtering it downstream?** Cleaning bad data inside the AI layer is expensive and never ends. The discipline that compounds is fixing it at the transactional system itself.

4. **Do your transactional, analytical and AI workloads share a foundation?** If they don't, every time an agent succeeds and you want to understand why, you are stitching systems together again. That overhead is permanent until the architecture changes.

Chapter 5

Governance Is Not the Brake. It's the Engine.

Most executives assume governance slows AI down. The opposite is true: the organizations moving fastest are the best-governed. Governance, done correctly, is not a constraint on speed — it's what makes speed safe enough to sustain. Most organizations get stuck in a cycle that David Meyer, SVP of Product at Databricks, describes as the trap of the agentic era — moving too fast with no governance, creating sprawl and risk, then slamming the brakes to install governance that was never designed to scale.

Both states are expensive. The organizations that escape this cycle will govern at the platform level rather than the process level — which means central teams hold full observability, model restrictions are handled at the foundation and individual teams experiment freely without waiting on review cycles where people without context have to decide "yes" or "no."

What changes when governance moves to the platform is the unit cost of trust. In the dashboard era, governance was a once-per-project negotiation: a security review, a compliance sign-off, a data access approval. In the agentic era, the volume of things to govern outpaces the human workflow that used to govern them. Agents spin up, call tools, write data and hand off tasks to other agents.

Every action is a governance event. Process-level governance cannot keep pace with that volume; only platform-level governance can. The organizations that recognize this stop treating governance as a phase of project delivery and start treating it as a property of the foundation everything runs on.

This is the shift Kunal Taneja, who leads field engineering for APJ at Databricks, sees in how application design itself is changing. In the past, governance was an afterthought — something you did to satisfy an IT audit or pass a regulator's review. But in a world of generative AI and agents, governance is almost always the first thing to start with. Building strong data operations, MLOps capability and a way to govern data as it moves into production are not compliance activities — they're the foundation that makes everything else trustworthy at scale.

The most actionable governance concept for organizations trying to move fast without creating risk is what Lexy Kassan, Lead Data and AI Strategist at Databricks, calls the "paved path." Once an architecture has been validated against the organization's security, risk and governance requirements, every new AI use case that follows it does not start from scratch. The technology has been proven, the security structures are in place and the risk mitigations are documented. This means the hurdles that would otherwise slow a project down can be removed without removing human oversight. Many use cases can follow the same path, and the marginal cost of governance per use case approaches zero. That is how governance becomes an accelerant.

Sal Companieh, Chief Digital and Information Officer at Cushman & Wakefield, has spent four years building exactly this kind of paved path across a 108-year-old commercial real estate firm with 53,000 colleagues worldwide. The measurable shift: time from idea to business outcome compressed from months to days. The mechanism is structural — a product operating model that embeds technologists in every business unit, a capital investment model that funds firm-wide capabilities regardless of which P&L benefits and enterprise standards that allow flexibility at the unit level without compromising the common platform. Governance, in that frame, is what allows the organization to move quickly and trust what it produces.

> "The enthusiasm for AI hasn't subsided. It has been well-rounded with the recognition that healthy, governed, scalable data is what actually accelerates outcomes."
>
> — Sal Companieh, Chief Digital and Information Officer, Cushman & Wakefield

Hanlin Tang, CTO of Neural Networks at Databricks, sees a dimension of governance that most organizations miss entirely. The conventional read on access controls is defensive — keep the agent away from salary data, away from regulated records, away from anything it should not touch. But the same controls that prevent exposure also tell the agent which slice of a sprawling lakehouse is actually relevant to the question being asked.

If five versions of a customer table sit in the catalog and the agent has access only to the most recent one, it cannot accidentally reason from yesterday's numbers. For agents, Hanlin argues, governance is not just security. It is also quality. Permissions are a signal of what to focus on and what to ignore — which is why governance belongs not next to agent quality, but inside the system that produces it.

The mental models that have served security teams for decades — test it, deploy it, patch it — do not apply to probabilistic AI systems. Omar Khawaja, Field CISO at Databricks, puts the security dimension that most organizations underestimate this way: just because an organization has been using AI for three years does not mean it understands how AI works. People have been driving cars for a hundred years without being able to explain how a carburetor works. Security teams now have to be willing to unlearn as much as they learn.

Every other technology managed by security teams was deterministic. This one is probabilistic. The old rules do not all apply. It's impractical to attempt to mitigate all the risks of AI — rather, security/risk teams must first identify which risks actually apply and matter most, then surgically mitigate them.

The security teams that recognize this, and adopt what Khawaja calls a “hyper-growth mindset” about what they need to learn and unlearn, are the ones enabling their organizations to move with speed and confidence. The key to doing this: don’t start with a list of all the controls — start with the risks and let them dictate the specific controls warranted.

Khawaja is also direct about where the threat environment stands. The mean time to exploit a vulnerability collapsed from 23 days in 2025 to 1.6 days in 2026. Attackers are already using AI agents to scan systems, discover vulnerabilities and execute coordinated attacks at machine speed, around the clock. Defenders who rely on manual workflows and fragmented architectures simply cannot keep pace.

The response can’t be a new security silo bolted onto everything else. It has to be an extension of the foundation the rest of the organization is already building — same platform, same governance model, same observability, with security as a property of the architecture rather than a parallel system. Which gets to the broader principle: an agent is like a data asset. It continuously costs money, it has a threat surface and if you don’t know where it is, you can’t turn it off. Whatever discipline you apply to data governance — lifecycle policies, expiry rules, lineage, observability — you should apply to agent governance too.

The four governance decisions to make before your next AI deployment:

1. **Define your paved path.** Document the security, risk and compliance requirements that every AI use case must meet. Do the work once. Apply it everywhere.

2. **Govern at the platform, not the project.** Build observability, model restrictions and access controls into the foundation everything runs on — not into individual project reviews. Process-level governance cannot keep pace with agent-era volume.

3. **Treat agents like data assets.** They continuously cost money, they have a threat surface and you cannot turn off what you cannot see. Apply the same lifecycle policies — expiry, classification, observability — that you already apply to data.

4. **Don’t attempt to mitigate all possible risks.** Pick the ones that actually apply and matter in your context.

Chapter 6

The Operating Model Is the Strategy

Technology decisions and organizational decisions are inseparable in the agentic era. The diagnostic questions that reveal how advanced an organization's AI posture really is are not technical ones — they're organizational ones. Dael Williamson, Field CTO of EMEA at Databricks, asks four of them:

1. How close is the data and AI leader to the CEO?
2. Are data strategy and AI strategy one conversation or two?
3. Does the organization have an accurate inventory of its data assets?
4. Does it have a genuine, expanded definition of data — one that includes images, Slack messages, code, email and everything else that carries intelligence about the business?

The answers to those four questions, more than any technical benchmark, predict whether an organization will compound its AI investment or dissipate it.

What Williamson sees in the organizations gaining real ground is that the people shaping AI strategy genuinely understand how the technology works — not in a surface-level way. The pace of change demands it. Something that didn't work six months ago may work today because of a new model or a new technique. And the leaders pulling ahead are the ones who can track that and act on it quickly. AI rewards organizations that treat it as a living capability rather than a finished product.

Beyond the four diagnostic questions, four characteristics consistently distinguish the organizations generating real value from those producing demos:

- They are not taking the easy path of buying AI from every SaaS vendor — that produces tool sprawl without enabling the cross-functional collaboration that creates enterprise-scale value
- They run their AI portfolio like a venture company, killing projects that are not working and doubling down on the ones that are
- They have reconsidered build versus buy, with insourcing of core AI capability making a clear comeback
- The data and AI function has a seat at the senior leadership table — not as a function IT reports to, but as a peer of the CEO

What ties those four characteristics together is an operating model: a deliberate concentration of standards, platforms and governance at the center, paired with deliberate flexibility for the business units that have to execute. Most organizations get one half of that equation right and the other half wrong

Either the center owns everything and innovation suffocates, or the units own everything and the platform fragments into something nobody can govern. The organizations compounding their AI investments have figured out how to do both simultaneously. And they treat the operating model that holds those two ideas in tension as the strategy itself, not as a layer underneath it.

The clearest example of that operating model in practice is at Albertsons Companies, the $80 billion food and drug retailer with roughly 2,300 stores. Sunil Gopinath, who leads data and AI for the company, calls it a franchise model.

The organization stopped tolerating fragmented business-unit experiments and committed to one team, one platform, one operating mode. They focused on four strategic bets in customer experience, merchandising intelligence, labor and supply chain.

The center owns common horizontal components: governance, security, a central repository of reusable models, ingestion templates, feature store patterns and model monitoring. Application teams build on top of that foundation.

In nine months, more than 90% of engineers are engaging with AI tools, 1.38 million lines of AI-generated code have been accepted into the codebase and merchants now ask complex questions in natural language through Genie and get governed answers without writing a query or filing a ticket.

> "It's a franchise model. Common infrastructure, standards and governance at the center; local execution and innovation at the edges. Reusable accelerators let any team plug in and move 10 times faster. The platform isn't there to constrain innovation — it's there to compound it."
>
> — Sunil Gopinath, Head of Data and AI, Albertsons Companies

There is a counterweight to all of this momentum that Naveen Zutshi, CIO of Databricks, presses on with his peers in every conversation. The temptation in the agentic era is to move fast on every front simultaneously — modernize the legacy stack, deploy agents into production, build the operating model, run governance to catch up later. The pressure is real, and it produces a predictable mistake: organizations that have just spent five years escaping SaaS sprawl reproduce the same problem with agent sprawl.

"Stop feeding the beast of legacy" is Naveen's framing. Start the modernization if it hasn't already started. But don't replace SaaS sprawl with agent sprawl — be deliberate about which agents go into production and why. And do not treat AI governance as an afterthought. There is enormous pressure to move fast, but deferred governance always costs more later than it saves now.

The operating model is not only a question of structure. It's also a question of what people inside that structure are actually being asked to do.

Robin Sutara, Americas Data and AI Enablement Leader at Databricks, believes most organizations are underestimating the workforce shift at the level of work itself. The mistake is treating AI as something employees should add to their existing responsibilities without removing anything.

The result is cognitive overload — people doing their old jobs and the AI version of their old jobs simultaneously. The organizations getting this right are creating genuine space, protected time and explicit mandates for employees to ask a different question entirely: if I could design this process from scratch with the AI capabilities I have today, what would I build?

That is the deeper shift Sutara sees underneath the operating model conversation. Leaders are no longer managing deliverables. They're managing systems and processes that deliver at scale — which is a cultural change as much as a structural one. The measure of work moves from how many things you produced to whether you have built something that can produce as many of those things as the business needs, because you taught it how.

Five questions to test your operating model for in the agentic era:

1. **How close is your data and AI leader to the CEO?** Proximity signals strategic importance and determines how fast decisions get made.
2. **Are data strategy and AI strategy one conversation or two?** If they're separate, they're working against each other — and the gap shows up everywhere downstream.
3. **Have you concentrated standards, platforms and governance at the center while leaving real flexibility at the edges?** Most organizations get one half of that equation right and the other wrong. The franchise model gets both right.
4. **Are you running your AI portfolio like a venture company?** Kill what's not working, double down on what is and refuse to defend projects out of organizational momentum.
5. **Are you redesigning work, or piling AI on top of it?** Adding AI capabilities to unchanged job descriptions produces cognitive overload. Designing the work with AI from the start produces compounding returns.

Chapter 7

Five Strategic Plays for the Next 18 Months

It is the unglamorous, laborious work of getting your data in order — the cataloging, the semantic definitions, the governance decisions that nobody puts on a slide — that creates the payoff. The organizations doing that work now are not just preparing for AI. They're building the only kind of advantage that cannot be copied quickly: a foundation that compounds.

01

Define the outcome before you design the system

The organizations generating real AI value do not start with a model or a platform. They start with a specific outcome they want to drive — a workflow they want to own, a decision they want to accelerate — and engineer backward from there. If you cannot describe what success looks like before you build, you're not ready to build.

02

Engineer your architecture to bring context to the model

A frontier model has no native sense of what your business means by "active customer," which version of a margin number was flagged for a one-time promotion or which policy applies to which region. The architecture is what closes that gap.

Get the data foundation right — catalog, semantic layer, governed access — before layering AI on top. Confirm that your transactional infrastructure can support agents that branch, experiment and recover at machine speed. Design for a unified platform rather than stitching together point solutions you will spend years integrating. Architecture is not the goal — context is the goal. Architecture is how you deliver it.

03

Rethink governance for agents, not just for data

The governance frameworks most organizations have in place were built for a deterministic world — structured data, human-run applications and predictable outputs. Agents are none of those things; they're probabilistic, parallel and fast.

Applying old governance thinking to new agentic systems doesn't slow you down safely — it creates the illusion of control while the real risks accumulate. Build your paved path. Govern agents the way you govern data assets. Make platform-level governance the prerequisite for organizational speed.

04

Redesign the work — don't just put AI "on top of it"

This is the dimension most organizations underestimate, and the one with the longest lead time. Enabling a workforce to do a fundamentally different kind of work is not a training initiative. It's a redesign of what work itself looks like in the organization, and it requires protected time, explicit mandates and leadership that understands what is actually being asked of people.

The shift is from managing deliverables to designing systems that deliver at scale. The organizations that resource this properly will capture the productivity gains. The ones that treat it as a communications exercise will not.

05

Stop feeding the beast of legacy while you build what comes next
Modernization and transformation are not the same project, but they have to run in parallel. Every resource tied up maintaining infrastructure that was not built for the agentic era is a resource not building toward it. Make deliberate decisions about what to wind down. Don't replace SaaS sprawl with agent sprawl — be intentional about what you build, what you buy and what you retire.

The clearest example of all five plays running simultaneously inside one organization is at Aer Lingus, the 90-year-old Irish flag carrier. About 18 months ago, the airline redirected a sizable portion of its IT and change capital away from legacy maintenance and toward a unified Databricks-based foundation — governance, data quality and what Dave O'Donovan, Chief Digital, Data and Transformation Officer, calls a Data Literacy Academy.

The discipline he runs the program on is a 75/25 split: roughly three-quarters of capacity on long-term foundational work, one-quarter on rapid wins that keep the business engaged. A dedicated continuous improvement team works across finance, customer care and operations to redefine processes so they're AI-ready before agents arrive. Dave's framing for what makes this kind of work durable is quite basic: don't try to be future-proof. The technology changes every 6 to 12 months. Be pivot-ready instead.

What You Build Now Determines What You Can Do Next

The executives in this eBook have moved past the question of whether AI will transform their industries. They're working on the harder question: what does it actually take to build something that lasts?

The answer that emerges across every conversation in these pages is the same. It's not a single technology decision. It's the accumulation of unglamorous, deliberate work — getting the data in order, governing the agents, redesigning the workflows, building the people capability — before anyone outside the organization can see why it mattered. That work does not show up on a demo. It shows up in the compounding returns of organizations that started earlier than everyone around them thought was necessary.

The executives in these pages stopped asking what AI can do — they started asking what they could build before anyone could see why it mattered. That is the work this book is about. And it's the only kind of work, in this market, that compounds.

The platform behind the decisions

The decisions in this eBook — unified governance, a semantic layer, transactional infrastructure for agents, AI that reasons on your own data — are the same four demands every architecture answers: context, control, cost and choice. The Databricks Data Intelligence Platform is designed so they compound together rather than competing for the same budget and team. Unity Catalog governs data, models and agents in one place — control and cost in a single layer.

Genie brings the context to reason over that governed data without rebuilding your semantic layer. Lakebase gives agents the transactional foundation to branch, experiment and recover at scale. Lakewatch extends that governance to security operations. Built on open formats, none of it locks you in — the choice stays yours. Agent Bricks brings production-grade agentic capability to organizations ready to move from experimenting to operating.

The architecture in this eBook is not theoretical. It's what our customers are building, and what the platform is designed to support.

About Databricks

Databricks is the data and AI company. More than 20,000 organizations worldwide — including adidas, AT&T, Bayer, Block, Mastercard, Rivian, Unilever and 70% of the Fortune 500 — rely on Databricks to build and scale data and AI apps, analytics and agents. Headquartered in San Francisco, with 30+ offices around the globe, Databricks offers a unified platform that includes Lakebase, Genie, Agent Bricks, Lakeflow, Lakehouse and Unity Catalog. To learn more, follow Databricks on LinkedIn, X, YouTube and Instagram.

Key Databricks capabilities referenced in this eBook

Unity Catalog: unified governance for data, models, agents and AI assets across all environments

Genie: conversational analytics powered by your company's own data, semantics and business terminology

Lakebase: transactional database built for the agentic era, integrated with the lakehouse platform

Lakewatch: open, agentic SIEM built on the security lakehouse, announced in March 2026

Databricks Apps: the application layer that brings AI-powered intelligence to business users

Agent Bricks: end-to-end tooling for building, evaluating and deploying AI agents at enterprise scale

To learn how to create an actionable roadmap for advancing your AI capabilities, download the Databricks AI Maturity Model.

databricks

Executive Contributors

The *New Architecture of Enterprise AI* is the result of candid conversations with Databricks executives and field leaders. Thank you to the following contributors, whose perspectives shaped every insight in this report.

Customer Executives

Neal Bradbury, Chief Product Officer, Barracuda
Sal Companieh, Chief Digital and Information Officer, Cushman & Wakefield
Dave O'Donovan, Chief Digital, Data and Transformation Officer, Aer Lingus
Stephen Ecker, Chief Data Officer, Trinity Industries
Sunil Gopinath, Head of Data and AI, Albertsons Companies
Kristy Mayer-Mejia, Global Head of Data Transformation, Kraken
Nader Mherabi, Chief Digital and Information Officer, NYU Langone Health

Databricks Executives

Arsalan Tavakoli-Shiraji, SVP, Field Engineering, Databricks
Naveen Zutshi, CIO, Databricks
Ari Kaplan, Global Head of Evangelism, Databricks
Craig Wiley, VP of AI, Databricks
Dael Williamson, Field CTO, EMEA, Databricks
David Hackett, VP, Field Engineering, Databricks
David Meyer, SVP, Product, Databricks
Hanlin Tang, CTO, Neural Networks, Databricks
Junta Nakai, VP, Financial Services, Databricks
Kunal Taneja, AVP, Field Engineering, APJ, Databricks
Lexy Kassan, Lead Data and AI Strategist, Databricks
Maria Zervou, Chief AI Officer, EMEA, Databricks
Nick Eayrs, VP, Field Engineering, APJ, Databricks
Omar Khawaja, Field CISO, Databricks
Raela Wang, Director of Field Engineering, Databricks
Robin Sutara, Americas Data and AI Enablement Leader, Databricks
Shanku Niyogi, VP, Product Management, Databricks

Databricks Content Leads

Josh Howard, Senior Director, Product Marketing, Databricks
Catherine Brown, Senior Product Marketing Manager, Databricks
Aly McGue, Senior Customer Marketing Manager, Databricks