BARC

# From Data to Agents

## Modernize Your Architecture for the Age of AI

Author: Kevin Petrie

Publication: September 2025


## Abstract

This research note explores how organizations should upgrade their architectures to succeed with agentic AI. We identify popular use cases, define must-have requirements, and explore the architectural underpinnings of effective data, model, and agentic workflows. We conclude by recommending specific steps to get started.


Research sponsored by:

Informatica

# Table of Contents

# Introduction

Generative AI (GenAI) is writing the next chapter in business reinvention. As with steam engines, electricity, and personal computing, this breakthrough offers a powerful opportunity for humans to create more value with less effort. Business leaders are justifiably excited: GenAI models now outperform human cognition on some levels and inference costs are falling faster than electricity or PC memory ever did. A rising cadre of early adopters boast about big boosts to efficiency. Journalists showcase bullish case studies across industries, focusing in particular on "agents" that make decisions and take actions based on GenAI outputs.

And yet, GenAI and agentic AI introduce unprecedented complexity and risk. A web of interrelated elements – storage systems, data pipelines, AI models, applications, and tools – add up to many points of failure. With their livelihoods and reputations on the line, many business leaders agree with the recent observation in Harvard Business Review that "organizations aren't ready for the risks of agentic AI."

***Agentic AI introduces unprecedented complexity and risk.***

History reassures us that, with persistence and caution, disruptive innovations are worth the initial failures. In 1879, inventor Thomas Edison botched some electrical wiring and nearly burned down William Henry Vanderbilt's house. But in subsequent years Edison, backed by Vanderbilt, architected safe power grids that drove the Industrial Revolution.

This research note seeks to answer the pressing question of our era: how can organizations upgrade their architectures to achieve similar economic gains with agentic AI? We identify popular use cases, define must-have requirements, and explore the architectural underpinnings of effective data, model, and agentic workflows. We conclude by recommending specific steps to get started.

## Agentic AI defined

Agentic AI is a sophisticated form of AI that makes decisions and takes action autonomously to achieve specific goals and outcomes in dynamic environments. Agents are AI-enabled applications designed to execute simple tasks, from simple transactions to navigating complex decision trees for smart actions and valuable outputs. An agent's "perception" of its environment and surrounding context allows it to work in a human-like fashion to solve problems and execute tasks while operating autonomously at scale and with limited oversight. AI agents can become smarter over time by learning to optimize their efficiency, representing the next step in insight-driven actions beyond simple rule-based, policy-driven robotic process automation.

***Agentic AI is a sophisticated form of AI that makes decisions and acts autonomously.***

All this requires a governed architecture that combines new and old elements across the storage, data, model, and agent layers. The problem is that many enterprises cannot yet do this. In fact, just 36% of organizations "strongly believe they have the right architecture and tools to enable data-driven decisions and support large volumes of data," according to a recent survey by Unisys. Their heritage SQL-based

databases, data warehouses and analytics tools were designed for a different era. They need new tools and techniques that handle agentic workflows in a governed fashion while integrating with established systems.

## Challenges

Many obstacles lie in the way of modernization. Data and IT teams have accumulated technical debt, making short-term project decisions that limit future options. They wrestle with siloed datasets, complex workflows, and manual processes. Many stakeholders – from data engineers to machine learning engineers, data scientists, and application developers – lack the necessary skills to optimize new technologies. Governance issues, data quality issues in particular, pose another significant challenge: BARC research finds that one third of data professionals do not trust the inputs they feed to AI models.

## Benefits

Teams that overcome these challenges can capture the compelling business benefits of agentic AI. They can enable AI innovation, especially to achieve the primary objectives of boosting efficiency and reducing cost. They can make their employees more productive, increasing output, reducing toil, and freeing up time to focus on strategic work. A modern architecture also sets the foundation for future innovative projects, eliminating friction for new technologies and processes. It helps organizations engage customers in creative ways and spawn new revenue streams. It also unlocks the value of proprietary datasets, creating sustainable competitive advantage.

***A modern AI architecture unlocks the value of proprietary datasets, creating sustainable competitive advantage.***

## Use cases

Popular emerging use cases for agentic AI include coding, customer service, e-commerce, supply chain management, finance, and marketing/sales. These agents, often packaged as commercial software products, employ GenAI and other analytical techniques to solve problems, automate tasks, and advise humans.

- **Coding**. Software development is by far the most popular use case for GenAI and agentic AI. Developers across industries, technology in particular, use commercial offerings such as Microsoft's Gitub Copilot to build and test all types of programs.
- **Customer service**. Chatbots help service managers or even customers answer questions and resolve issues. Customer service is a leading use case for agentic AI, according to a recent survey by a16z.
- **E-commerce**. Chatbots also guide website visitors as they peruse, evaluate, and buy products. More than 20% of CIOs now have GenAI in production to recommend customer purchases, according to a16z.

- **Supply chain management**. Logistics managers use agentic AI to analyze and optimize supply chains. For example, Uber Freight has 30 agents that route and track shipments across its network.
- **Finance**. Accountants, procurement managers, and other finance personnel automate processes, for example by drafting and reviewing contracts. BNY uses agents to accelerate trade settlements by proactively addressing risks.
- **Marketing/Sales**. GenAI models and agents generate promotional content, devise marketing programs, and identify sales opportunities. Many marketing teams already use GenAI to draft content, and companies such as Walmart are rolling out agents to help their teams and partners prepare formal outreach programs.

Leading adopters often implement "teams" of agents to tackle several functions at once. Walmart addresses four of the use cases described here.

# Architecture

## Must-have characteristics

To support such use cases, a modern architecture must be integrated, governed, secure, modular, open, flexible, efficient, and AI-assisted.

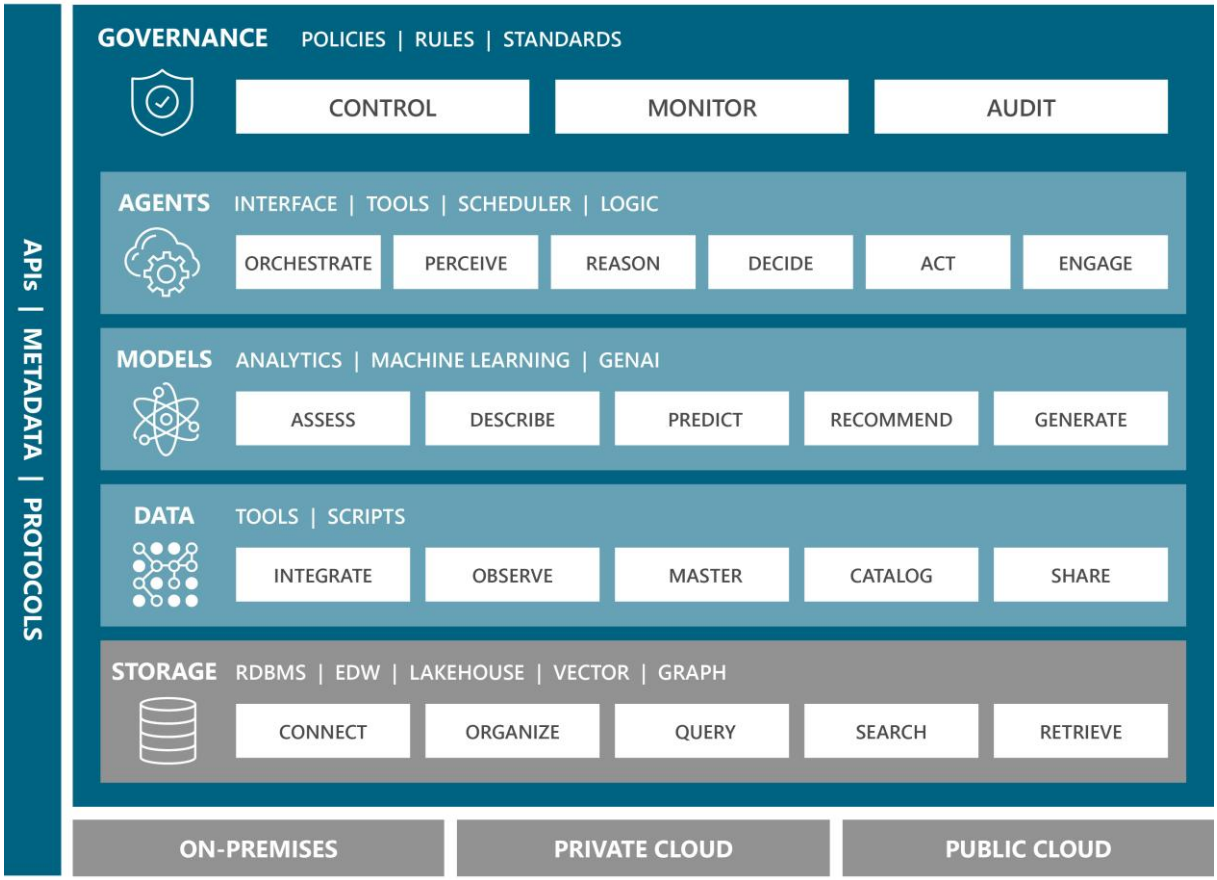
- **Integrated**. A modern architecture integrates data, models, and applications across infrastructure that resides on-premises and in the cloud. New commercial products can help by combining models and applications into a single workflow.
- **Governed**. A modern architecture governs these elements to control risks and ensure trustworthy inputs, outputs, and outcomes.
- **Secure**. The corollary to governance is security. Organizations must protect their systems from penetration, usage, or disruption by bad actors.
- **Modular**. As with Lego bricks, architectures need standard, interchangeable, and interoperable building blocks – data pipelines, models, scripts, etc. – that they can assemble and reuse as needed.
- **Open**. A modern architecture also needs open connectivity with today's vibrant ecosystem, including its various tools, libraries, formats, and APIs. This enables teams to share data and metadata across departments and organizations.
- **Flexible**. Modular and open systems provide organizations with the flexibility they need to expand, shrink, or reconfigure workflows when business requirements change.
- **Efficient**. A modern architecture must consume cloud resources and AI tokens efficiently while giving teams full visibility and control over usage costs.
- **AI-assisted**. AI assistants make data engineers, data scientists, ML engineers, and developers more productive. They help design, build, test, debug, deploy, and optimize elements across the enterprise – and enable less technical users to contribute as well.

***A modern architecture must be integrated, governed, secure, modular, open, flexible, efficient, and AI-assisted.***

## Architectural elements

A modern architecture has four layers: storage, data, models, and agents. It governs these layers with shared tools and glues everything together with application programming interfaces (APIs), metadata, and protocols. Let's work from the bottom up in our diagram, highlighting where and how to upgrade for agentic AI.

**Modern Agentic Architecture**




### Storage Layer


A variety of platforms store data for analytics and AI. They connect to and ingest diverse operational datasets, then query, search, and retrieve them for consumption by models and applications. The methods vary by platform type. Relational database management systems (RDBMS) organize rows and columns in tables to support operational workloads, while data warehouses (EDWs) organize tables for analytics. Lakehouses merge the concepts of EDWs and data lakes to manage tables alongside semi-structured objects such as telemetry logs or unstructured objects such as documents, emails, and

images. Vector databases, meanwhile, organize documents or other high-dimension data as numerical embeddings. And graph databases capture the inter-relationships of concepts and entities.

**Modernization**. New converged platforms help modernize data storage for agentic AI. Many database and lakehouse vendors, from gorillas like Google Spanner to startups like Vespa, now combine storage types to reduce administrative effort and consolidate multi-source, multi-format data for AI. However, enterprise environments remain complex: most organizations have multiple RDBMS and EDW systems along with one or more lakehouses, vector DBs, or graph DBs. To support agentic AI workflows, data and IT teams need comprehensive, consistent methods of managing these distributed datasets across data centers, private clouds, and public clouds.

***New converged platforms modernize data storage for agentic AI.***

## Data Layer


This brings us to the data layer. Data architects, engineers, and stewards manage distributed datasets with commercial tools and homegrown scripts. They integrate data using a variety of pipeline patterns. Traditional ETL pipelines extract tabular data from an RDBMS, then transform and load it into an EDW or lakehouse, and recent variations change the sequence to ELT or ELTL. Data teams use observability tools to detect, assess, and resolve data quality issues, and they use master data management (MDM) tools to maintain "golden records" that standardize business terminology. Vendors such as Informatica offer suites that consolidate many such capabilities into a suite, helping deliver high-quality inputs to AI models.

**Modernization**. Data teams are adopting new tools and scripts to manage documents, images, and other unstructured objects that contain valuable context for agentic AI. They use modern catalogs to centralize metadata for these diverse data objects as well as the models and agents that layer on top. To broaden the possibilities for agentic AI, they use marketplaces to share data, models, and agents with other teams and organizations.

***Data teams must adapt their tools to manage unstructured objects that contain valuable context for agentic AI.***

The lifecycle of data operations (DataOps) helps data teams optimize data management for agentic AI. It applies the developer operations (DevOps) practices of testing and continuous (CI) integration and continuous delivery (CD) to the creation, deployment, and iteration of data pipelines. This lifecycle helps data engineers learn to build effective pipelines for unstructured data. They can make frequent changes as they label images, convert text to vector embeddings, and so on.

## Model Layer


Now we come to the models that consume all this data for training and inference. To start, basic analytical models such as a linear regression assess and describe the relationship between variables: price, revenue, and so on. Various machine learning (ML) models, available through public libraries such as PyTorch or TensorFlow, go further to predict or recommend outcomes. Now GenAI models such as ChatGPT from OpenAI and Claude from Anthropic generate content in the form of text, imagery, or even audio and video.

**Modernization**. AI teams have different model types contribute to agentic AI workflows, each according to their specialty. Data scientists might train one ML model to predict customer actions and another to recommend purchases based on those predictions. A GenAI model, meanwhile, might generate text that delivers this recommendation to the customer through a mobile chatbot. Retrieval-augmented generation (RAG) workflows operate in the background, retrieving relevant and contextual data and injecting it into the GenAI prompt to make outputs more accurate.

***Various specialized models contribute to agentic AI workflows.***

The lifecycle of model operations (ModelOps) helps AI teams optimize the models that contribute to agentic AI. ModelOps enables frequent adjustments to the creation, deployment, and iteration of ML models in particular. This lifecycle helps data scientists train and build effective ML algorithms, and it helps ML engineers deploy and iterate them in response to changing business conditions. These stakeholders also can use ModelOps to implement and refine ML models that evaluate the accuracy, toxicity, and relevance of GenAI outputs. Another variation of ModelOps, also called LLMOps, optimizes the operation of GenAI models within these agentic workflows.

## Agent Layer

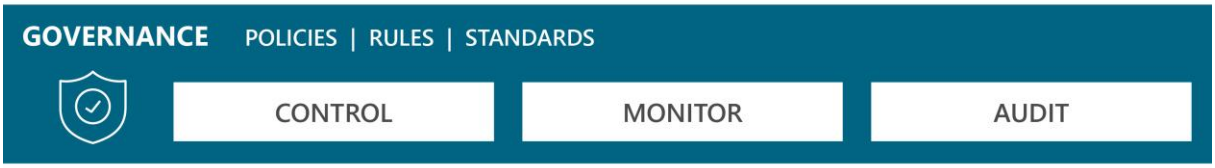

Agents, the culmination of a modern architecture, tie everything together and bring AI to life. Their logic receives human questions or instructions through a chatbot interface, then assigns tasks to various tools. One tool might have a GenAI model interpret instructions, and other tools might execute tasks or query GenAI/ML models to respond to the instructions. The agent uses a scheduler to orchestrate sequences of tasks among tools and models.

***Agents tie everything together and bring AI to life.***

Together these elements perceive situations, analyze options, and make decisions. Agents then act on their decisions and even engage external applications or agents as they do so. Most agents so far remain semi-autonomous and require human oversight to stay within safe guardrails.

The lifecycle of agent operations (AgentOps), as the name suggests, helps AI teams build, deploy, and iterate agents. Data scientists and ML engineers must test, monitor, and tune complex agentic workflows, especially their interactions with GenAI and ML models. This requires close inspection of agent outputs and actions, with a focus on safety.

## Governance




Governance programs enforce policies, rules, and standards to reduce risks of a modern architecture. The policies define overarching objectives and principles, the rules define how to enforce them, and the standards provide technical structure. Data teams must collaborate with AI teams and developers to organize these elements across the data, model, and agent layers, holding stakeholders accountable and flexibly adapting to new or changed requirements.

### *Governance policies, rules and standards reduce the risk of agentic AI*

- **Data**. Data can misrepresent the truth, breach privacy, exhibit bias, and compromise intellectual property. Traditional governance tools mitigate these risks with controls for data observability, usage authorization, and cataloging. They also monitor and audit activities in these areas to comply with internal policies and external regulations. Organizations now must extend their data governance programs to address the risks of models and agents.

- **Models**. Models can have opaque, confusing logic or generate toxic outputs that treat humans unfairly. Modern governance programs must reduce these risks with explainable logic and safe outputs. To ensure explainability, AI teams adopt new lineage tools, model registries, catalogs, and techniques such as SHAP or LIME that quantify feature contributions to model outputs. To ensure safe outputs, they implement evaluator models and content filters. Such measures control risk, monitor operations, and audit compliance.

- **Agents**. Agents can draw incorrect or inappropriate conclusions, act in damaging ways, or even try to subvert human intentions. Modern governance programs must ensure safe decisions, controllable actions, and permissible behavior. To mitigate such risks, AI and developer teams implement overrides, kill switches, and other controls that respond to alerts from evaluator models. They also might trigger human interventions or double-checking by alternative GenAI models.

APIs, metadata, and protocols

**APIs | METADATA | PROTOCOLS**

So how do all these moving parts get along with one another? Application programming interfaces (APIs), metadata, and protocols fill this role within a modern agentic architecture.

- **APIs** expose functionality across each layer, enabling agents to interact with data stores, invoke models, and trigger downstream applications. They standardize how agents call services, smoothing the integration of enterprise systems and tools to make architectures more open and flexible.
- **Metadata** adds essential context that agents need to operate intelligently. It describes the structure, lineage, classification, sensitivity, and quality of data, while also recording performance and purpose of models and agents. Modern catalogs should capture all such metadata and make it readily available for DataOps, ModelOps, and AgentOps as part of agentic AI projects.
- **Protocols** coordinate interactions among system components, ensuring agents can understand and trust what they receive. The Model Context Protocol (MCP) standardizes access to data from source platforms and applications, helping RAG workflows ensure models generate grounded responses. The Agent-to-Agent (A2A) protocol allows agents to share tasks, decisions, and status updates as part of a collaborative workflow.

***APIs, metadata, and protocols help agentic AI elements interoperate with one another.***

# Guiding Principles for Modernization

Agentic AI marks the next leap in enterprise automation, reducing the need for humans to make operational decisions and take action. To capitalize on this opportunity, organizations must modernize their architecture across four critical layers: storage, data, models, and agents. They must integrate traditional and modern technologies while governing each layer and optimizing the DataOps, ModelOps, and AgentOps lifecycles. Organizations that modernize their architectures effectively can unlock productivity, reduce costs, and gain sustainable competitive advantage.

Data, AI, and business leaders should take the following steps to get started.

- **Start with a high-impact, low-risk use case**. Identify a business bottleneck – such as content generation, customer service, or logistics – and scope how to ease the problem with incremental changes to existing processes. By tackling a small, solvable problem, your teams will gain confidence and gather operational feedback without introducing excessive risk. If the project succeeds, they can win executive support and budget for broader modernization efforts.

- **Assess your architecture and pinpoint the gaps**. Inventory your current data, model, and application infrastructure to understand which systems support agentic workflows and which require upgrades. Pay particular attention to gaps in unstructured data management, model orchestration, and agent integration. This assessment will guide your team how best to modernize, one element at a time.

- **Take a lifecycle approach to governance**. Form a working group that includes data engineers, data scientists, ML engineers, developers, and compliance stakeholders. Task them with defining policies, rules, and standards across the DataOps, ModelOps, and AgentOps lifecycles. This team should review use cases, evaluate risks, and implement oversight mechanisms across each architectural layer.

- **Train your teams to work with new tools and workflows**. Skills shortages are a major obstacle to AI success. Upskill your staff in areas such as prompt engineering, GenAI safety, vector search, and unstructured data processing. Encourage experimentation with AI-assisted development and automated pipeline tools to improve productivity. Also consider partnering with vendors or consultants to bridge initial knowledge gaps and accelerate deployment.

- **Consolidate and automate functionality**. Broad-based commercial products such as Informatica simplify data management, governance, and other architectural layers by combining functions within automated workflows. Evaluate how such products might boost your team's productivity, improve scalability, and shorten time to production.

- **Build for modularity, interoperability, and reuse**. Adopt modern tools and platforms – preferably broad commercial products – that support open APIs, metadata catalogs, and protocols such as MCP and A2A. Also look for modular pipelines, models, and agents that can be assembled and reassembled with minimal friction. Such building blocks will enable you to evolve from isolated use cases to collaborative agent teams across departments.

By taking these steps, organizations can responsibly harness the power of agentic AI. They can reduce risk, drive innovation, and create new business value in the age of autonomy.

# About BARC

BARC is the leading analyst firm for data & analytics, AI, corporate performance management (CPM) and ESG with a reputation for unbiased and trusted advice. Our expert analysts deliver a wide range of research, events and consulting services for the data & analytics community. Our innovative research evaluates software, vendors and service providers rigorously and highlights market trends, delivering insights that enable our customers to innovate with data, analytics and AI. BARC's 25 years of experience with data strategy & culture, data architecture, organization and software selection helps clients transform into truly data-driven organizations.

**Research**

BARC user surveys, software evaluations and analyst advisory services along with expert driven content such as research notes, trend analysis and blogs give organizations the confidence to make the right decisions. Our independent research gets to the heart of market developments, evaluates software, vendors and service providers thoroughly and gives valuable ideas on how to turn data, analytics and AI into added value and successfully transform businesses.

**Consulting**

The BARC consulting practice is entirely focused on translating companies' requirements into future-proof decisions. The holistic advice we provide helps companies successfully implement their data & analytics strategy and culture as well as their architecture and technology.

BARC's research and experience-founded expert input sets organizations on the road to the successful use of data & analytics, from strategy to optimized data-driven business processes.

**Events**

At BARC events, leading minds and industry experts come together to share insights and drive innovation. Our conferences, roundtables and online webinars attract over 10,000 participants annually, offering a unique blend of information, inspiration and interactivity. These events provide a platform to exchange ideas with peers, explore emerging trends and gain expert perspectives on market developments.

By engaging with thought leaders and industry practitioners, participants discover actionable strategies to enhance their business and stay ahead in the evolving world of data & analytics.

BARC

# About Informatica

Informatica (NYSE: INFA), a leader in AI-powered enterprise cloud data management, helps businesses unlock the full value of their data and AI. As data grows in complexity and volume, Informatica's Intelligent Data Management Cloud™ delivers a complete, end-to-end platform with a suite of industry-leading, integrated solutions to connect, manage and unify data across any cloud, hybrid or multi-cloud environment. Powered by CLAIRE® AI, Informatica's platform integrates natively with all major cloud providers, data warehouses and analytics tools— giving organizations the freedom of choice, avoiding vendor lock-in and delivering better ROI by enabling access to governed data, simplifying operations and scaling with confidence.

Trusted by approximately 5,000 customers in nearly 100 countries—including over 80 of the Fortune 100—Informatica is the backbone of platform-agnostic, cloud data-driven transformation. Informatica. Where data and AI come to life.™

**Contact info**

**Informatica Headquarters**

2100 Seaport Blvd
Redwood City, CA 94063
USA
Find an office near you

informatica.com
Contact us

**Data Decisions. Built on BARC.**

**www.barc.com**

**Germany**
BARC GmbH
Berliner Platz 7
D-97080 Würzburg
+49 931 880651-0

**Austria**
BARC GmbH
Hirschstettner Straße 19 / I / IS314
A-1220 Wien
+43 660 6366870

**Switzerland**
BARC Switzerland GmbH
Buchhaldenstraße 7
CH-5442 Fislisbach
+41 76 340 35 16

**United States**
BARC US
13463 Falls Drive
Broomfield, CO 80020
USA