

## SOLUTION BRIEF

# EDB Postgres® AI Agent Factory

Solving the 10 critical barriers to enterprise AI at scale



## Executive summary

Enterprises are rapidly investing in AI factories—platforms designed to build, deploy, and scale AI agents and generative AI applications. However, the majority of these initiatives fail to reach production or deliver sustained value. The reason is not a lack of models or data but the emergence of a new class of system-level challenges that traditional data and application architectures were never designed to handle.

AI systems are fundamentally different. They are probabilistic, stateful, multicomponent pipelines that require continuous coordination between models, data, memory, and external tools. What appears to be a simple agent interaction often masks a highly complex system involving retrieval pipelines, orchestration layers, model routing, and governance controls. As a result, organizations encounter a consistent set of failure points—technical issues that translate directly into business risk. These issues define the gap between AI experimentation and production-grade AI systems.

EDB Postgres AI (EDB PG AI) Agent Factory is designed specifically to close this gap by addressing the **10 critical issues that prevent AI factories from scaling**.

## The reality of AI factories: Why most fail before production

The fundamental barrier to scaling AI in the enterprise is not model capability. It is, instead, the accumulation of systemic challenges that emerge when AI moves from experimentation to production. Across industries, organizations encounter the same 10 issues: inconsistent agent memory, degrading retrieval quality, stale data pipelines, unreliable tool orchestration, security vulnerabilities, uncontrolled model costs, lack of determinism, poor multi-agent coordination, limited observability, and gaps in governance. These are not isolated technical problems; they are structural weaknesses that compound across AI systems, leading to unreliable outputs, operational risk, and stalled deployments.

EDB PG AI addresses these challenges holistically by embedding solutions directly into a unified, Postgres-native platform and ensuring that every layer of the AI system, from data and memory to orchestration and governance, is consistent, secure, and production ready from the start.

# The 10 critical issues in AI factories

At the core of every failed or stalled AI initiative are 10 systemic challenges:

1. The primary issue is **state management and agent memory**. AI agents must maintain consistent context across interactions, yet most architectures rely on fragmented memory systems. Without synchronized persistence strategies, agents operate on stale or conflicting context, leading to incorrect or inconsistent outputs.
2. Closely related is **retrieval quality and vector drift**. AI systems depend on embeddings and vector search to retrieve relevant context, but, over time, embeddings degrade as data changes. Without proper versioning and reindexing strategies, retrieval accuracy declines and agent reasoning suffers.
3. **Data freshness and pipeline latency** represent another critical challenge. AI agents require access to current, consistent data, but pipelines often introduce latency or fail to synchronize with source systems. This results in outdated outputs that undermine trust in the system.
4. **Tool invocation and orchestration** introduce additional complexity. AI agents frequently interact with external systems, APIs, and workflows. Without robust retry logic, idempotency controls, and failure handling, these interactions can break, producing incomplete or hallucinated results.
5. Security risks emerge through **prompt and policy injection**, in which malicious or unintended inputs manipulate agent behavior. Without runtime enforcement and input isolation, agents can execute unsafe instructions.
6. **Model routing and cost control** become critical as organizations deploy multiple models across different providers. Without intelligent routing and fallback strategies, costs escalate and performance becomes unpredictable.
7. The probabilistic nature of AI introduces challenges in **determinism and testing**. Traditional testing frameworks struggle to validate nondeterministic systems, making debugging and validation difficult.
8. As organizations scale, **multi-agent coordination** becomes a bottleneck. Agents must collaborate, resolve conflicts, and avoid redundant work, yet poorly designed systems lead to deadlocks or infinite loops.
9. **Observability and debugging** are equally challenging. AI pipelines span multiple components, and without end-to-end tracing, failures become opaque and difficult to diagnose.
10. Finally, **governance and data security** remain foundational. AI systems must enforce strict controls over data access, embeddings, and outputs, particularly in regulated industries. Without embedded governance, organizations risk data leakage and compliance violations.

## The EDB PG AI solution: A unified, Postgres-native platform

EDB PG AI addresses these challenges through a single, unified Postgres-native platform for building, testing, and deploying AI agents. It combines a visual agent development environment, automated data pipelines, native vector search, and flexible model serving.

By colocating data, models, and orchestration, EDB PG AI eliminates the need for external systems and the integration complexity they create. This unification delivers three connected advantages: Development and production become the same environment, so teams ship directly without reengineering; all AI data stays within Postgres, so existing governance extends to AI workloads automatically; and a visual development environment puts production-grade agents within reach of every development team, not just AI specialists.

## Business impact: From AI experimentation to production scale

EDB PG AI transforms AI from a fragmented experimentation effort into a scalable, production-ready capability.

First, it accelerates time to value. Organizations can move from concept to production in weeks rather than months by eliminating the need for separate development and production environments. McKnight Consulting Group found that EDB PG AI enables delivery [3x faster than DIY cloud](#), compressing time to production from 28 weeks to just 9 weeks.

Second, it reduces complexity and cost. By consolidating multiple systems into a single platform, EDB PG AI eliminates integration overhead and reduces operational burden. McKnight Consulting Group found 67% lower development effort and 38% lower maintenance complexity compared to DIY cloud approaches, enabling smaller teams to deliver larger outcomes.

Third, it enhances security and governance. By keeping all AI data—vectors, embeddings, and knowledge bases—within the Postgres environment, EDB PG AI ensures that existing governance controls extend to AI workloads. This eliminates the need for new security boundaries and reduces compliance risk.

Finally, it enables an agentic workforce—not a future state but a pattern already emerging in advanced enterprise deployments. Agents in production fall into three categories: operational agents that handle high-volume tasks and surface only what requires human judgment; generalist agents that give every employee a reliable path to enterprise data without IT overhead or sovereignty



risk; and knowledge agents that keep institutional expertise current and queryable. These categories are where the 10 systemic challenges above become most consequential. Retrieval drift degrades knowledge agents. Orchestration failures stop operational agents mid-workflow. Governance gaps make generalist agents a compliance liability. EDB PG AI addresses all of this at the platform level, so the capacity these agents return is real and sustained, not a pilot that stalls before it scales.

## Solving the 10 issues: Platform-level integration

The key advantage of EDB PG AI is that it does not address these challenges individually—it resolves them at the platform level.

State management is unified within PostgreSQL, ensuring consistent agent memory. Retrieval quality is maintained through integrated vector engines and indexing strategies. Data pipelines are automated and synchronized with source systems, ensuring freshness.

Tool orchestration is embedded within the platform, enabling reliable execution of workflows. Security is enforced through runtime policies and data governance controls. Model routing is managed dynamically, optimizing cost and performance.

Testing and determinism are improved through versioning of models, prompts, and data. Multi-agent coordination is handled within a unified orchestration framework, preventing conflicts and inefficiencies. Observability is integrated end to end, providing visibility into system behavior.

Governance is enforced at every layer, ensuring compliance and data protection.

## Comparative analysis

Organizations typically approach AI development through three models: DIY architectures, cloud-based AI services, or platform-based solutions.

### EDB PG AI vs. Alternative Approaches

	DIY AI stack (LangChain, vector DBs, APIs)	Cloud AI services	Point solutions (vector DBs, orchestration tools)	EDB PG AI
<b>Strengths</b>	Flexibility, customization	Fast setup, managed infrastructure	Optimized for specific functions	Unified platform, sovereign control, integrated stack
<b>Limitations</b>	High complexity, integration overhead	Vendor lock-in, data movement, security concerns	Fragmentation, integration challenges	Requires Postgres as the data foundation
<b>Business impact</b>	Slow time to market, high cost	Limited control, compliance risk	Partial value, high complexity	Fastest path to production, lowest risk

\*Competitive comparisons are based on publicly available information and are subject to change as vendor offerings evolve and new information is made available. All product names, trademarks, and registered trademarks are the property of their respective owners.

### Technical Capability Comparison

Capability	EDB PG AI	DIY stack	Cloud AI services	Point solutions
<b>Data integration</b>	Native in Postgres	External pipelines	Cloud managed	Separate tools
<b>Vector search</b>	Native	External database	Managed service	Separate system
<b>Model serving</b>	500+ providers, flexible	Manual integration	Vendor limited	Partial
<b>Orchestration</b>	Integrated	Custom built	Limited	Fragmented
<b>Governance</b>	Built in	Manual	Vendor controlled	Partial
<b>Deployment model</b>	Anywhere (on prem, cloud, hybrid)	Complex	Cloud only	Mixed

\*Competitive comparisons are based on publicly available information and are subject to change as vendor offerings evolve and new information is made available. All product names, trademarks, and registered trademarks are the property of their respective owners.

# Why EDB PG AI wins

EDB PG AI succeeds because it addresses the root cause of AI failure: fragmentation. Instead of requiring organizations to assemble and manage multiple systems, it provides a unified platform on which all components are integrated and optimized to work together.

It eliminates the prototype-to-production gap by unifying development and runtime environments. It removes data movement by keeping all AI workloads within Postgres. It reduces complexity by embedding orchestration, observability, and governance into the platform.

Most important, it enables organizations to build and deploy AI systems that are not only functional but reliable, scalable, and secure.

The following table maps each of the 10 critical issues to EDB PG AI solutions at both the technical and business levels.

## EDB PG AI: Big 10 challenges → Platform solutions

Technical issue	Core problem	EDB PG AI solution	Business outcome
<b>State management and agent memory</b>	Inconsistent agent context across interactions	Postgres-native persistence with unified relational and vector storage, ensuring synchronized, durable agent memory across sessions	Reliable, context-aware agents that behave consistently across workflows
<b>Retrieval quality and vector drift</b>	Declining accuracy of embeddings over time	Native vector engine with versioning, reindexing, and high-performance retrieval at scale	Sustained accuracy in AI reasoning and decision-making
<b>Data freshness and pipeline latency</b>	Stale or delayed data that agents operate on	Integrated data pipelines with direct access to operational Postgres data—no external sync required	Real-time, trusted outputs aligned with the current business state
<b>Tool invocation and orchestration</b>	Unreliable execution of external tools and workflows	Built-in agent orchestration with Langflow-based visual pipelines, retry logic, and API-native execution	Reliable automation and reduced workflow failure rates
<b>Prompt/policy injection</b>	Exposure to unsafe or malicious instructions	Postgres-governed data + runtime policy enforcement + no external AI boundary	Secure AI execution with enterprise-grade protection
<b>Model routing and cost control</b>	Inefficient use of models, leading to cost overruns	Flexible model serving (500+ providers) with dynamic routing, fallback strategies, and cost-aware selection	Optimized AI spend with predictable performance
<b>Determinism and testing</b>	Difficulty validating probabilistic systems	Full versioning of prompts, models, data, and pipelines within a unified environment	Reproducible AI behavior and faster debugging cycles
<b>Multi-agent coordination</b>	Conflicts and inefficiencies between agents	Native orchestration framework for agent collaboration, task control, and loop prevention	Scalable multi-agent systems without deadlocks or redundancy
<b>Observability and debugging</b>	Lack of visibility across complex pipelines	End-to-end observability with pipeline tracing, token-level insights, and performance monitoring	Faster root cause analysis and reduced downtime
<b>Governance and data security</b>	Risk of data leakage and compliance violations	Sovereign AI architecture—data, vectors, and models remain inside Postgres with full policy enforcement	Compliance-ready AI with no new security boundary

## What this means strategically

The table above highlights a critical distinction: Most AI platforms address the challenges individually and reactively, requiring additional tools, integrations, and operational overhead. In contrast, EDB PG AI addresses them systemically and natively.

Because all components—data, vector search, orchestration, model serving, and governance—are integrated into a single Postgres-native environment:

- There is **no data movement**, eliminating latency and security risks.
- There are **no disconnected systems**, eliminating integration complexity.
- There is **no replatforming requirement**, preserving existing investments.
- There is **no new operational model**, reducing adoption friction.

This is what enables EDB PG AI to compress time to production from months to weeks and, per Everest Group research, [reduce overall complexity by more than 55%](#) while maintaining full enterprise control.

The “Big 10” challenges are not edge cases—they are the fundamental reasons most AI initiatives fail to scale. EDB PG AI succeeds because it eliminates these challenges at the architectural level, not as afterthoughts. Instead of assembling an AI stack, organizations adopt a platform on which:

- Memory is consistent.
- Retrieval is accurate.
- Data is current.
- Workflows are reliable.
- Security is enforced.
- Costs are controlled.
- Systems are testable.
- Agents coordinate effectively.
- Behavior is observable.
- Governance is built in.

This is the difference between experimenting with AI and **operating AI at enterprise scale**.

## EDB Postgres AI: The sovereign data and AI platform for the agentic enterprise

EDB PG AI brings together a unified data layer, governance, sovereign control and orchestration, and an agent runtime environment, giving enterprises a trusted foundation for AI on infrastructure they own and control. The platform unifies transactional, analytical, and AI workloads in a single Postgres-based architecture—eliminating ETL, data movement, and operational fragmentation. And you choose where and how to deploy: on-premises, cloud, managed, or certified appliance.

The outcome: production-ready sovereign AI in days or weeks, not months.



EDB Postgres® AI (EDB PG AI) is the sovereign data and AI platform for the agentic enterprise. Built on Postgres, the world's leading open source database, EDB PG AI unifies transactional, analytical, and AI workloads in a single governed architecture, on-premises and across clouds. To learn more, visit [www.enterprisedb.com](http://www.enterprisedb.com).

© EnterpriseDB Corporation 2026. All rights reserved.