Buy, boost, or build: Rethinking the AI decision in 2026

Five surprising realities that will help you decide whether to buy an AI system or build one

As the hype around artificial intelligence (AI) begins to subside, it is time for a more critical evaluation of the technology's actual impact on business outcomes.

This article distils key, often counter-intuitive, insights from recent research to help you navigate the critical 'build versus buy' decision for generative and agentic AI tools. Much is at stake: According to an MIT study, 95% of generative AI pilots fail to deliver measurable P&L impact, and making the right build-versus-buy decision is a key determinant of success.

Key Recommendations

- Buy where mature vendor solutions exist: external partnerships succeed about twice as often as internal builds.
- Use a hybrid approach – buy, boost, build – to balance speed, customisation and differentiation
- Treat production agents as factories: guardrails, orchestration, continuous RAG maintenance and testing are essential.
- Stop shadow AI by tightening governance, discovery and approved platforms.
- Shift focus from cost to speed-to-learning: validate value quickly and iterate

Takeaway #1: 'Buying' succeeds twice as often as 'building'

This result may surprise leaders who prefer to build and own core technology (often drawing on open source models), but the data is clear. MIT NANDA's 2025 *GenAI Divide* study found that "external partnerships see twice the success rate of internal builds."

Vendor solutions offer prebuilt integrations, specialised expertise, and business-focused metrics that link performance to ROI. In contrast, internal projects often become bogged down in technical complexity before delivering value. A 2025 S&P Global survey reported that 42% of companies abandoned their AI projects, often due to cost overruns and data and security concerns.

Takeaway #2: It's not enough to build a tool: you need to build an entire factory

Creating a production-ready AI agent is not a standalone software project. It requires building and maintaining a complete supporting infrastructure, often overlooked in initial planning. This infrastructure includes several complex components that must be designed from the ground up:

- **Multiple guardrails:** An AI agent without robust safety measures is a liability. A production system requires a three-layer defence framework to manage threats at every stage. The *input layer* acts as a perimeter, using scanners to redact sensitive data and threat detectors to block malicious prompts, such as indirect prompt injection. The *action layer* validates the agent's intended plan against business policies (e.g., trade value caps) and triggers human-in-the-loop approvals for high-risk operations. Finally, the *output layer* uses an LLM-as-judge pattern (a separate LLM validating the system's output) for hallucination detection and runs compliance checks to prevent violations of industry regulations. (OWASP Top 10 for LLM Applications, 2025)
- **Orchestration and observability infrastructure:** Agentic systems are defined by their ability to decide their own actions dynamically. This requires a sophisticated architectural layer to manage the thought-action-observation (TAO) cycle, in which the agent reasons, executes a tool, and learns from the results. As there is a trade-off between predictability and adaptability, a robust orchestration engine is needed to manage the stateful, non-deterministic control flow. Furthermore, debugging failures requires a specialised *agent observability* stack to trace the entire graph of TAO cycles, which is fundamentally different from traditional software logging that only catches code exceptions. (Yao et al., 2022)

- **Continuous RAG maintenance:** The retrieval-augmented generation (RAG) pipeline that connects an agent to your company's data requires ongoing maintenance. Without vigilance, the accuracy of a custom-built agent will decline over time.
- **Evaluation and testing platforms:** As AI agents are non-deterministic, traditional software testing is insufficient. A formal evaluation framework is needed to rigorously test for critical failure modes, such as tool selection accuracy, parameter correctness, and error handling.

Takeaway #3: The buy or build decision isn't a binary one

Enterprise leaders are moving beyond a simple binary choice, to consider a spectrum of options: buy, boost, or build. This strategic, hybrid approach can evolve as your company's AI maturity and business needs evolve.

Path	Description	Key benefits	Key drawbacks
Buy	Adopting an off-the-shelf, ready-to-use AI solution from a vendor.	Speed and simplicity; quick adoption without needing to invest in development.	Limited competitive differentiation; dependence on the vendor's roadmap and pricing.
Boost	Enhancing a vendor's solution by using proprietary or company-specific data.	More accurate and relevant results; customisation without starting from scratch.	Increased usage costs and prompt lengths; requires strong data governance.
Build	Developing, running and maintaining a custom AI solution from the ground up, often using open-source models.	Complete control and potential for significant competitive differentiation.	Expensive, difficult and requires significant upfront investment and advanced capabilities.

A flexible strategy is key: You might buy a solution for one department, enhance a tool with proprietary data for another, and reserve a full build for a core, differentiating capability.

As you develop your AI strategy, consider not only what you *can* build, but also what you can realistically own, maintain, and secure over time.

Takeaway #4: The biggest threat to your AI strategy is shadow AI

Organisations need to clamp down on the unsanctioned use of AI solutions, known as "shadow AI." Shadow AI accounts for most enterprise AI use, according to recent analysis (Netskope Threat Labs, 2025). If unaddressed, this trend undermines any centralised build-or-buy strategy, as employees are already making these decisions independently.

Shadow AI is proliferating across the enterprise in several distinct forms:

- **Cloud-based apps:** With employees frequently using personal, unmanaged accounts with generative AI apps for business purposes, the average organisation now uses seven different cloud-based AI applications. Generative AI data violations more than doubled in 2025, according to Netskope Threat Labs.
- **GenAI platforms:** Employees are using powerful platforms, such as Azure OpenAI, Amazon Bedrock, and Google Vertex AI, to build custom applications that interact directly with enterprise data stores.
- **On-premises tools:** A growing trend is the use of on-premises tools, such as Ollama, and agent frameworks, such as LangChain. These deployments are often the most difficult for central IT to discover and secure.

Widespread usage of shadow AI creates a large, unmonitored attack surface, increases the risk of sensitive data leaks, and fragments the organisation's efforts, preventing it from realising the full value of its AI investments.

Takeaway #5: The core criteria has changed from 'cost' to 'speed-to-learning'

AI has fundamentally altered the strategic calculus for software decisions. Historically, the build versus buy decision was framed by development time and budget. Today, the most critical constraint is the time it takes to validate whether a new AI capability delivers real, measurable business value. In an AI context, speed is no longer about *how fast can we ship a feature*, but *how fast can we learn whether this capability is a real generator of value or a demo*.

Buying or enhancing an existing solution enables organisations to reach the "learning and validation" stage quickly. In contrast, a full internal build may take 12 months or more to match commercial options (RAND, 2024). By then, the market and competitors will have advanced. A shift from building to learning reflects the need to move on from AI hype to tangible business outcomes.

Conclusion: Asking the right questions for the agentic age

In 2026, boards and investors will expect enterprises to realise real value from AI. But the build-or-buy decision for AI solutions shouldn't be rushed. While building in-house may offer control, it carries hidden risks, higher failure rates, and longer time-to-value (per Gartner). Leaders often underestimate the architectural and operational complexity required to run an AI agent in production.

Agent maturity varies by domain and task: some are ready for a buy decision, others are not. If mature solutions that fit your requirements exist, consider them. If not and the business case is clear today, build pragmatically to capture value. Otherwise, consider a strategic 'wait' decision to leverage today's quick AI development pace.

Contacts

For further information, or to discuss your AI challenges, please contact us.

Jorg Schalekamp

Partner Engineering, AI and Data & AI Leader

JSchalekamp@deloitte.nl

Bram Lentz

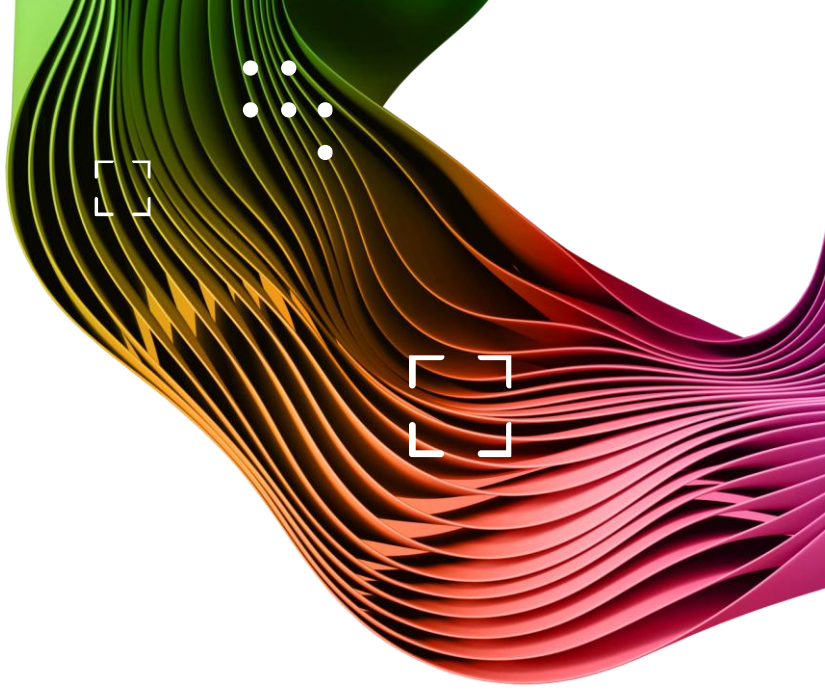
Partner

BLentz@deloitte.nl

Jeroen Smeets

Product Director

JeSmeets@deloitte.nl



Deloitte.

CoRe Creative Services. RITM2402012