

In recent years, the brewing industry has faced mounting pressure to modernize and optimize its operations. Among the many innovations driving this forward, artificial intelligence (AI) stands out—promising efficiency gains and smarter decision-making. But premature excitement often meets a hard wall: enterprise AI adoption stalls, not for lack of performance, but because of trust issues.

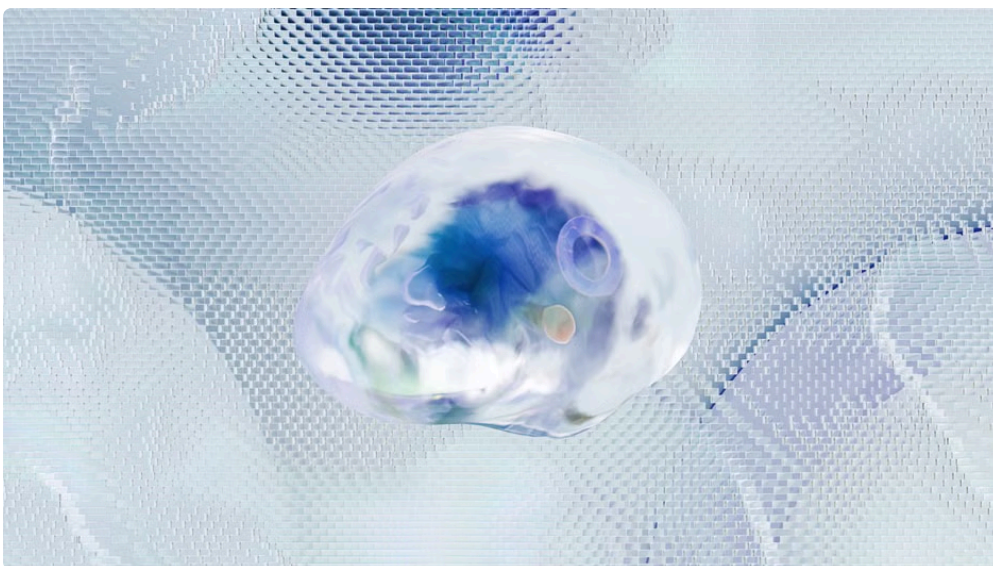
One brewer recently boasted a remarkable 20% higher rate of touchless planning—a leap in automation where the planning system executes with minimal human intervention. How did they achieve that? And why has this milestone proven elusive for others?

Let's explore the story through the lens of leading companies like Garrett Motion, PMI, and Pandora, and examine the vital roles that Large Language Models (LLMs) and retrieval-augmented generation (RAG) techniques played. We'll focus on the real bottleneck: trust. Along the way, we'll dissect key themes that often get overlooked, including explainability, consistency, business-logic alignment, and the dangers of hallucinations in AI outputs.

Trust: The Invisible Bottleneck in Enterprise AI Adoption

"Accuracy gains do not automatically create trust." This statement may sound counterintuitive, but it reflects what many enterprises discover the hard way. At face value, improved AI accuracy should translate to stronger confidence from planners, managers, and executives, enabling them to cede control and adopt touchless planning. Yet, while models may indeed perform better in benchmark tests or initial pilot runs, the real-world adoption curve remains steep when humans start interacting with AI outputs.

Garrett Motion, an automotive technology company, experienced this firsthand. Their AI-driven supply chain optimization models delivered superior results on paper, yet users hesitated to trust the model's suggestions without deep validation. This dynamic is typical — trust is a social and procedural construct far more than a numerical one.



- **Explainability matters:** Users want to understand *why* the AI made a particular recommendation, not just that it's probably right.
- **Consistency is a must:** Output variance—even from the same input—breeds doubt.
- **Alignment with business logic:** The model must reflect existing domain knowledge and constraints, not work at cross-purpose.

Without these factors addressed, accuracy improvements alone won't break through the "user skepticism" barrier inhibiting touchless adoption.

Explainability, Consistency, and Business-Logic Alignment: The Triad for Trust

The brewer who cracked the code to 20% higher touchless planning invested heavily in explainability frameworks integrated tightly with their AI tools. Here's how these three pillars came into play:

Explainability Through Traceability and RAG

Using retrieval-augmented generation (RAG) techniques, the planning system could pull in relevant, historically validated data snippets during its reasoning process, making the AI's output more transparent. Rather than a black box answer, planners saw the pieces of information the model retrieved to justify recommendations. This approach mirrors the strategy adopted by Pandora in their AI personalization platform, where RAG enhances content relevance and explainability.

This detail is critical for "variance review cycles," wherein planners compare AI outputs across different runs or input scenarios. The more the AI can elucidate its decision-making, the easier it becomes to trust its consistency.

Consistency Through Stable Prompting and Feedback Loop Planning

One notorious problem with vanilla LLMs is random output variance—even when given identical prompts. This unpredictability instills doubt and forces humans back into the loop for "sanity checks." To tackle this, the brewer imposed strict prompt engineering combined with ongoing feedback loop planning. The feedback loop ensured continuous evaluation and retraining of the model as domain conditions evolved, reducing variance and maintaining output reliability over time.

PMI, a global leader in consumer products, uses similar continuous improvement feedback loops to refine AI-powered forecasting models, achieving better consistency and lower error margins.

Business-Logic Alignment Through Integrated Rule Systems

Data-driven AI models sometimes produce recommendations that violate embedded business rules or operational constraints. The brewer adapted a hybrid approach where rule-based systems embedded in the AI pipeline acted as checkpoints, ensuring no output violated critical policies.

This alignment is crucial. When outputs contradict known business logic, trust inversely collapses, regardless of reported accuracy. Garrett Motion's experience validates this: their AI model initially generated supply chain plans that ignored supplier capacity limits, causing planners to distrust subsequent outputs until this integration improved.

Why Accuracy Gains Alone Don't Translate to Trust: The Hallucination Problem

One of the most pernicious issues AI planners face is hallucinations—situations where the model confidently fabricates information that appears plausible but is factually incorrect or contextually irrelevant. This "fabricated certainty" can wreak havoc when not caught early, leading to costly mistakes.

LLMs are especially vulnerable to this problem. Breweries, with their complex ingredient sourcing, seasonal <https://o9solutions.com/articles/can-you-really-trust-your-ai> demand fluctuations, and multi-stage production

processes, require high fidelity in planning inputs. A hallucinated forecast or capacity assumption can cascade into serious production disruptions.

Here's why hallucinations undermine trust regardless of accuracy scores:



- **Invisible Errors:** Not all hallucinations are obvious. When AI invents plausible but incorrect data, planners may miss these critical mistakes.
- **Disrupts Variance Review Cycle:** Inconsistent hallucinations across cycles confuse users about the model's reliability.
- **Feeds Black Box Excuses:** When flagged, AI vendors often revert to "the model decided," which erodes confidence further.

The brewer addressed hallucinations by integrating RAG-enabled retrieval steps that verified and cross-checked generated planning details against authoritative databases. This significantly reduced fabricated outputs and enhanced the credibility of each AI recommendation.

Continuous Improvement: The Feedback Loop Planning Advantage

Higher touchless planning rates are not a one-off achievement but a product of continuous improvement. The brewer's success hinged on a robust feedback loop planning process:

1. **Monitor Variance Review Cycles:** Regularly analyzing differences between AI outputs and actual outcomes pinpoints drift and inconsistencies.
2. **Incorporate User Feedback:** Planners annotate and provide corrections that feed back into model retraining and prompt tuning.
3. **Iterate Rapidly:** Frequent update cycles reduce model obsolescence and maintain alignment with evolving realities.

Garrett Motion's and PMI's deployments reinforce this theme: the AI system cannot remain static post-go-live. Instead, it needs a living process that blends human insights and technical refinements to sustain trust and performance.

What Would I Show in an Audit?

Let's question the claim: "20% higher touchless planning rates." To believe it, I'd want to see concrete audit artifacts demonstrating trust, consistency, and error management:

Audit Aspect Details to Provide Explainability Logs of RAG retrieval paths per recommendation, showing cited data sources and reasoning steps. Variance Review Cycle Metrics Historical records of AI output stability across multiple runs for identical inputs. Business-Logic Compliance Reports illustrating how each AI output adheres to rule-based constraints without violations. Hallucination Detection Instances of flagged fabricated outputs with root-cause analysis and remediation logs. Feedback Loop Implementation Change logs from model retraining cycles based on planner corrections and performance trends.

These artifacts move the discussion beyond black-box excuses to an accountable governance framework—precisely what enterprises need to scale touchless planning with confidence.

Conclusion

Trust, not technology alone, remains the pivotal bottleneck in enterprise AI adoption for supply chain and production planning. The brewer achieving 20% higher touchless planning rates cracked this trust code by integrating explainability via retrieval-augmented generation, reducing variance through feedback loop planning, and ensuring alignment with core business logic.

Garrett Motion, PMI, and Pandora spotlight how diverse industries converge on similar solutions: mitigating hallucinations, committing to continuous improvement, and refusing to accept accuracy as an end in itself.

Enterprises looking to improve their automated planning must focus on the human-machine partnership foundation. Metrics like touchless rates matter because they reflect collective trust—not just model performance. When AI recommendations are transparent, consistent, and aligned with reality, they become a true partner in planning, unlocking the full potential of feedback loop-driven continuous improvement.