

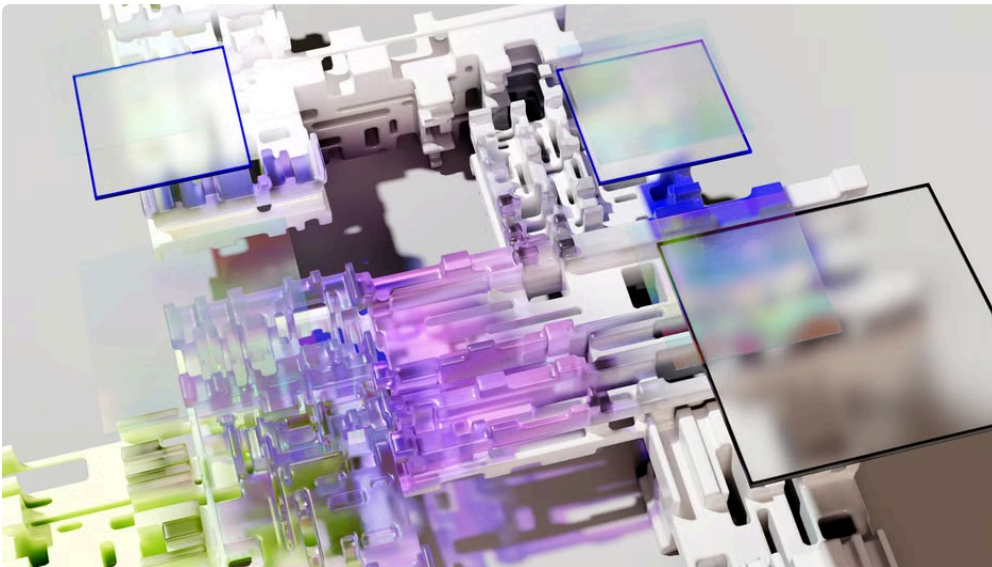
In the race to harness the power of Large Language Models (LLMs) for enterprise forecasting, accuracy often takes center stage. Companies like **Garrett Motion**, **PMI**, and **Pandora** are eagerly incorporating LLMs and advanced techniques like *retrieval-augmented generation (RAG)* to enhance predictive capabilities. But is improved accuracy enough to earn the trust needed for widespread adoption? Spoiler: No.

Why Trust is the Real Bottleneck in Enterprise AI Adoption

In enterprise environments, trust is the critical bottleneck for AI adoption—not raw accuracy numbers. Sure, LLMs can spit out impressive forecast accuracy metrics, but business leaders remain skeptical. Trust isn't built on numbers alone.

- **Garrett Motion** learned this the hard way, when despite strong accuracy improvements, teams hesitated to rely on LLM-generated forecasts without clear explanations.
- **PMI** invested heavily in AI models but continually encountered resistance due to concerns about inconsistent results and unclear decision logic.
- **Pandora** found that even with solid accuracy gains, governance frameworks needed bolstering to satisfy compliance and risk management stakeholders.

When forecasting directly impacts supply chains, finances, and strategic decisions, businesses require more than just numbers—they demand transparency, repeatability, and alignment with established business logic.



Accuracy vs Trust: What's Missing?

Accuracy is a necessary condition, but far from sufficient. Consider what trust actually entails:

1. **Explainability:** Can you break down how the forecast was generated? What data and rules influenced the model's predictions?
2. **Consistency:** Will the model produce stable results for identical inputs over time, or is there unpredictable variability?
3. **Business-Logic Alignment:** Does the model respect known constraints and domain knowledge, or does it sometimes make inexplicable leaps?

Most LLMs are infamous for their black-box nature. The phrase “the model decided” rings alarm bells for finance and operations teams. Without detailed audit trails, how do you prove that an aggregate prediction wasn’t pulled out of thin air?

Explainability: What Would I Show in an Audit?

For companies aiming to replace traditional forecasting tools with LLMs, the question is simple: *What would I show in an audit?* If compliance officers or regulators call, can you demonstrate the step-by-step reasoning linking input data, intermediate inferences, and final forecasts?



Retrieval-augmented generation (RAG) holds promise here, as it integrates external knowledge explicitly into generation. Instead of vague “training data memorization,” RAG models ground their answers in verifiable sources. Yet even then, traceability requires robust tooling and governance frameworks.

Hallucinations and the Risk of Fabricated Certainty

One of the most insidious risks when trusting LLMs for forecasting is hallucination—the confident generation of plausible yet false information.

- Imagine an LLM forecasting supply chain needs but fabricating demand spikes or ignoring real constraints.
- Such “fabricated certainty” is dangerous because it can mislead decision-makers who assume output reflects reality.

Evidence from enterprise pilots shows hallucinations can persist despite improvements in raw accuracy metrics. This disparity highlights why a sole focus on numbers can breed risk rather than confidence.

Consistency and Business-Logic Alignment: The Overlooked Metrics

Enterprise forecasting doesn’t operate in a vacuum. Forecast outputs intersect with legal contracts, vendor agreements, inventory capacities, and financial targets. This complex ecosystem demands models that:

- Consistently honor business rules and constraints
- Produce stable forecasts over repeated runs with the same data

- Clearly signal uncertainty and potential exceptions

Garrett Motion's experience underscores that gains in accuracy must be matched by rigorous validation of model consistency and alignment with domain knowledge.

Forecast Governance: Beyond Accuracy Metrics

To manage **LLM forecasting risk** effectively, companies need to embed these AI tools within strong governance frameworks. Such frameworks should encompass:

Governance Aspect	Description	Why It Matters
Auditability	Detailed logging of input data, intermediate steps, and rationale behind forecast outputs	Supports compliance and accountability; enables post-hoc analysis
Explainability Tools	Visualization and narrative explanations that translate model decisions into business terms	Builds stakeholder trust; aids interpretation and actionability
Consistency Checks	Automated tests that flag significant variability under stable conditions	Prevents erratic outputs; signals model degradation or bugs
Alignment Verification	Periodic validation that forecasts respect business rules, constraints, and domain logic	Ensures operational feasibility; reduces errors
Governance Policies	Defined roles, responsibilities, and approval processes for forecast adoption and overrides	Clarifies ownership; manages risk and accountability

For **PMI** and **Pandora**, embedding robust governance was crucial to turning promising LLM pilots into trusted operational systems.

Conclusion: Trust is the Outcome, Not the Input

In summary, relying solely on improved accuracy puts enterprises at risk. The journey to trusted LLM-driven forecasting demands tackling these hard challenges:

- Developing explainability that lives up to audit requirements
- Ensuring consistency and stability in forecast outputs
- Aligning forecasts with concrete business logic and domain expertise
- Mitigating hallucinations with retrieval-augmented generation and rigorous validation
- Embedding AI tools within comprehensive forecast governance frameworks

Companies like **Garrett Motion**, **PMI**, and **Pandora** exemplify the balanced approach needed—pursuing accuracy gains meanwhile investing heavily in transparency, reliability, and governance.

Without trust, no amount of accuracy can guarantee adoption. So next time you see an LLM forecast with <https://o9solutions.com/articles/can-you-really-trust-your-ai> dazzling accuracy claims, ask yourself: *what would I show in an audit?* If you can't answer that clearly, trust is still out of reach.