

Artificial Intelligence (AI) adoption in enterprises has surged, powered by advances in Large Language Models (LLMs) and retrieval-augmented generation (RAG). Yet despite impressive accuracy improvements, many organizations remain wary. The bottleneck? **Trust**. Without trust, no amount of technical promise will lead to widespread operational use.

Three companies—Garrett Motion, PMI, and Pandora—highlight the real-world challenges and stakes in deploying AI. Their stories underline an essential distinction: *explainability* versus *after-the-fact rationalization*. This difference gets to the heart of whether AI can become a reliable, auditable partner in decision-making or just an inscrutable “black box” spouting plausible-sounding stories.

## Trust: The Single Biggest Barrier to Enterprise AI Adoption

Organizations have been dazzled by the accuracy gains LLMs and RAG tools bring—better predictions, refined suggestions, and faster insight generation. But as Garrett Motion discovered in applying AI to manufacturing data, “accuracy” alone isn't enough. Their engineers needed **true traceability** in the AI's output to ensure every step could be audited and verified against physical realities.

Similarly, PMI's marketing analytics team found that initial AI models had volatile outputs, with the same inputs sometimes producing inconsistent results. Pandora's content curators faced hallucinations—completely fabricated facts presented with confidence—which threatened brand trust.

These experiences point to a core lesson: Trust emerges from consistent, step-by-step reasoning aligned with business logic—not simply blind belief in high accuracy scores.

## Explainability vs. After-the-Fact Rationalization

| Aspect                        | Explainability                                                                                 | After-the-Fact Rationalization                                                                              |
|-------------------------------|------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|
| Definition                    | Transparent, stepwise reasoning tracing output directly back to input data and business rules. | Post hoc justifications or plausible stories generated to explain AI decisions without real causal linkage. |
| Trust Impact                  | Builds confidence via auditability and true traceability.                                      | Erodes trust due to potential fabrication and inconsistent explanations.                                    |
| Alignment with Business Logic | Explicitly aligns AI decisions with company policies and domain expertise.                     | Tends to ignore or gloss over actual business rules, prioritizing convenience.                              |
| Handling Hallucinations       | Minimizes hallucinations by grounding in verifiable data.                                      | Often generates fabricated certainty that misleads stakeholders.                                            |
| Use Case Suitability          | Critical in compliance-heavy, safety-critical, or mission-critical environments.               | Can suffice for exploratory or creative tasks where accuracy is less scrutinized.                           |

## Why Explainability Matters

Explainability is not just a compliance checkbox. For Garrett Motion, integrating models that demonstrated every decision step enabled engineers to correct data issues and refine workflows. PMI's analytics workflows depended on audit trails that linked recommendations back to verifiable marketing inputs. Pandora's editorial teams mandated explainable AI to prevent spreading misleading information to listeners.

Explainability secures business-logic alignment, meaning AI outputs resonate with domain knowledge and organizational rules. Models following this discipline capture the “why” behind decisions—not just the “what.” This makes outputs predictable, consistent, and defensible.

## The Pitfalls of After-the-Fact Rationalization

After-the-fact rationalization masquerades as explanation but is often just clever storytelling. It's easy to generate persuasive-text "reasons" after an AI model outputs a prediction—but these reasons can be completely fabricated. RAG tools, while powerful, can fall into this trap when retrieval sources are incomplete or noisy, generating confident but false assertions.



In PMI's early adoption phase, rationalizations disguised inconsistent model logic, causing analysts to chase phantom trends. Pandora found that hallucinated facts risked audience trust, forcing a costly retrain of models with stricter grounding mechanisms.

## Why Accuracy Gains Alone Don't Build Trust

You might think the [o9solutions.com](https://o9solutions.com) obvious solution is: "If the model is more accurate, just trust it more." But true trust requires more than good numbers on historical test sets:

- **Consistency Matters:** Garrett Motion needed outputs that could be reliably reproduced with the same inputs every time. Random output variance breeds skepticism.
- **Transparency in Decision Paths:** PMI auditors demanded visible, justifiable links between data points and final recommendations, not just scorecards.
- **Resistance to Hallucinations:** Pandora insisted on explicit mechanisms to detect and prevent fabricated, confident-sounding claims. Hallucinations are trust killers.

Accuracy without these qualities feels like a black box—impressive, maybe even useful in some cases, but ultimately too risky for critical enterprise processes.

## How to Achieve True Traceability and Step-by-Step Reasoning

What does it take to build AI systems that offer the kind of meaningful, actionable explainability these enterprises require?

1. **Data Provenance Integration:** Maintaining strict linkages from input data to every intermediate step and output decision.
2. **Explicit Business Logic Encoding:** Embedding domain rules and policies into the model or its post-processing to ensure outputs respect organizational constraints.

3. **Hybrid AI Architectures:** Combining LLMs with symbolic reasoning or rule-based engines that produce inspectable reasoning chains.
4. **Retrieval-Augmented Generation with Auditable Sources:** Using RAG with curated, versioned knowledge bases that can be referenced and validated in explanations.
5. **Robust Model Monitoring:** Detecting output drift, inconsistencies, and hallucinations promptly, with feedback loops for re-training.

## Garrett Motion's Approach

Garrett Motion integrated AI within their manufacturing line by ensuring every AI prediction tied back to sensor data logs, maintenance records, and engineering standards. This true traceability let them run audits showing exactly how the AI arrived at any warning or alert.

## PMI's Analytics Transparency

For marketing decisions, PMI layered explainable AI techniques on top of LLM outputs, forcing the model's recommendations to include references to original campaign data and verified statistics. This step-by-step reasoning turned previously mystifying AI outputs into transparent, actionable guidance.



## Pandora's Hallucination Challenge

Using RAG, Pandora ensured that every content recommendation or fact inserted by AI linked directly to authoritative sources in real time. This architecture minimized fabricated certainty, preserving user trust and editorial integrity.

## Conclusion: Building Enterprise AI That Earns Trust

Enterprises face a clear choice: embrace real explainability or settle for after-the-fact rationalization. The former demands investments in tools, processes, and architectures that prioritize auditability, step-by-step reasoning, and true traceability. The latter risks creating a fragile house of cards—results that look good on charts but crumble under scrutiny.

Garrett Motion, PMI, and Pandora show that adopting AI responsibly means moving beyond “the model decided” hand-waving. Instead, it involves demanding demonstrable, transparent linkages between data, logic, and output—because that’s what builds trust, the single most important factor in unlocking the full potential of AI in enterprises.

*What would I show in an audit?* Every enterprise AI team should keep that question front and center.