BEFORE AI : WHY DATA QUALITY MATTERS FOR ENTERPRISES

BY: K. SUPRIYA

MARCH 2026

TABLE OF CONTENTS

- AI Everywhere: Hype vs Reality
- Why Innovation Still Stalls
- When AI meets Reality
- The Hidden Bottleneck: Data Quality
- When Data Misleads: The Faulty GPS
- What do we mean by Data Quality?
- Data Without Context is Useless: The suitcase without a tag- A Metadata Story
- The Metadata Blindspot
- What is Metadata (and Why it Matters)?
- What "AI-ready Data" Really Means?
- From Data to Innovation: What Must Change

AI EVERYWHERE : HYPE VS REALITY

AI AVALANCHE IN MARCH 2026

12+ AI Models in One Week: The March Avalanche Is Here

⚡AI Search⚡
@aisearchio · Follow

What a crazy week in AI! 🚀

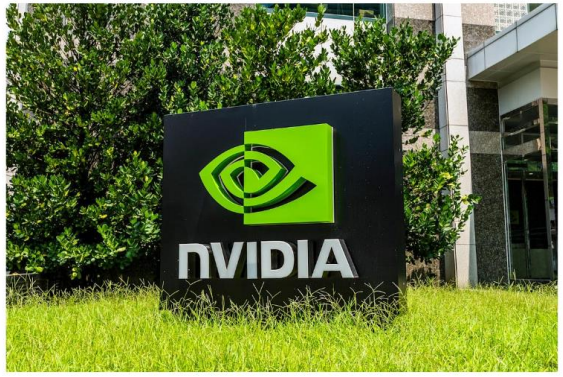
LTX 2.3
GPT 5.4
FireRed Edit 1.1
Kiwi Edit
HY WU
Qwen 3.5 small
Cuda Agent
CubeComposer
Helios
Spatial T2I
Spectrum
Utonia
& more!

EY has announced the launch of FlexiGenAI, an enterprise agentic AI platform.

Amazon News

Alexa+ launches in Canada, the first country to get the next generation of Alexa outside the US

Nvidia Is Building Its Own Open-Source Agent Platform



GEN-AI AND AGENTIC AI EXPLOSION

2026, ~40% of enterprise applications will embed AI agents (Source: McKinsey Report)

At Uber, AI agents are already generating ~8% of all code changes autonomously — up from near zero. (Source: Business Insider - March 2026)

~70% of North American enterprises are actively using AI (Source: NVIDIA Blog)

AI agents are expected to drive $2.6–$4.4 trillion in annual business value across use cases (Source: Gartner 2026)

Only 6% of companies fully trust AI agents to handle core business processes today.(Source: HBR (Survey 25-26))

Less than half (44%) say their data is adequate for AI, and only 39% have shared, usable data foundations. (Source: GenAI and Agentic AI Insights - Adobe)

Only ~34% of organizations are using AI to truly transform the business — most are still in experimentation mode(Source: Deloitte 2026 AI report)

Many organizations are still using GenAI as individual productivity tools rather than enterprise systems(Source: MIT Sloan 2026)

WHY INNOVATION STILL STALLS

The Broken Elevator

Building directory:

B2 – Data Quality: Out of order

B1 – Data Governance: "We'll circle back"

G – Basic Reporting: Flickering lights

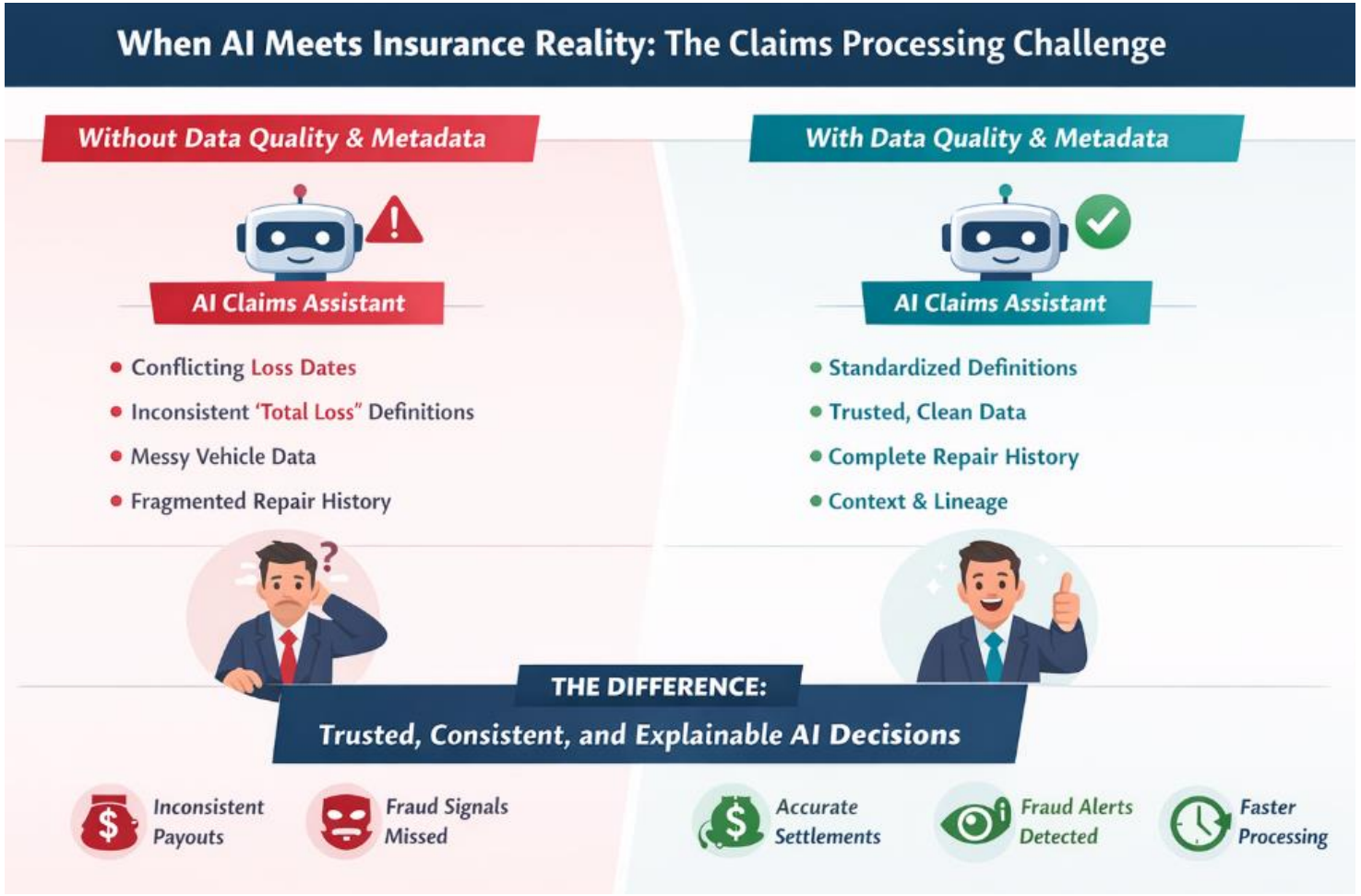
1 – Self-Service Analytics: Where most teams live

2 – AI Strategy: Where leadership wants to live

3 – Generative AI: Where the budget already moved

PH – Data-Driven Culture: Where the keynote promises we'll be in 18 months

WHEN AI MEETS INSURANCE REALITY



THE REAL BOTTLENECK: DATA QUALITY

- Data quality is the single most cited barrier to value realization from AI."
  Source: State of AI 2025: McKinsey Report

- At least 50% of generative AI projects were abandoned after proof of concept due to poor data quality, inadequate risk controls, escalating costs, or unclear business value."
  Source: Gartner (2026)

- Data, metadata, and governance are now treated as core enablers—not support functions
  Source: Gartner (2026)

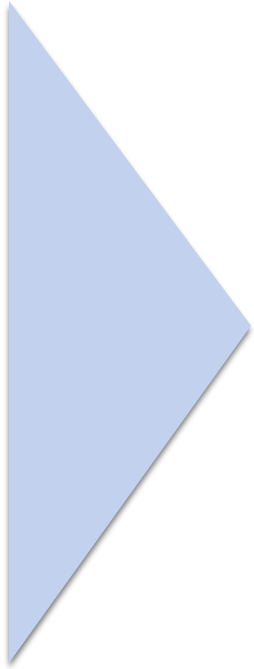
- Poor data quality costs companies in all industries an average of ~$12.9 million per year.
  Source: (Gartner)

WHEN DATA MISLEADS: THE FAULTY GPS

GPS Problem	DQ Dimension	
Sends you to the wrong location	Accuracy	*The data is incorrect. It's confidently wrong.*
No street names or landmarks	Completeness	*Key fields or reference data are missing.*
Same address gives different results each time	Consistency	*Different systems interpret or format the data differently.*
Directions are 3 hours out of date	Timeliness	*The data is stale — a closed road is still shown as open.*
Shows in miles when you expected kilometers	Validity / Format	*The data isn't in the expected format or unit.*
No idea how the route was calculated	Lineage /Traceability	*You can't trust what you can't trace — no metadata or audit trail.*
App crashes halfway there	Usability	*Even good data in a bad tool is useless.*

WHAT DO WE MEAN BY DATA QUALITY?

- It is the data's suitability for a user's defined purpose.
- Data quality is dependent on context and on the needs of the data consumer.



- Accuracy
- Timeliness
- Completeness
- Consistency
- Integrity
- Conformity/Validity
- Uniqueness

THE BIG CONCEPTUAL SHIFT

Traditional Data Quality Management	Automated AI-Ready Data Quality Pipelines
Classic approach used in many enterprises.	Modern approach used in AI-driven systems
Reactive	Proactive
Ownership resides with Data Stewards	Ownership resides with Data, ML and Platform teams

DATA WITHOUT CONTEXT IS USELESS: THE SUITCASE WITHOUT A TAG - A METADATA STORY

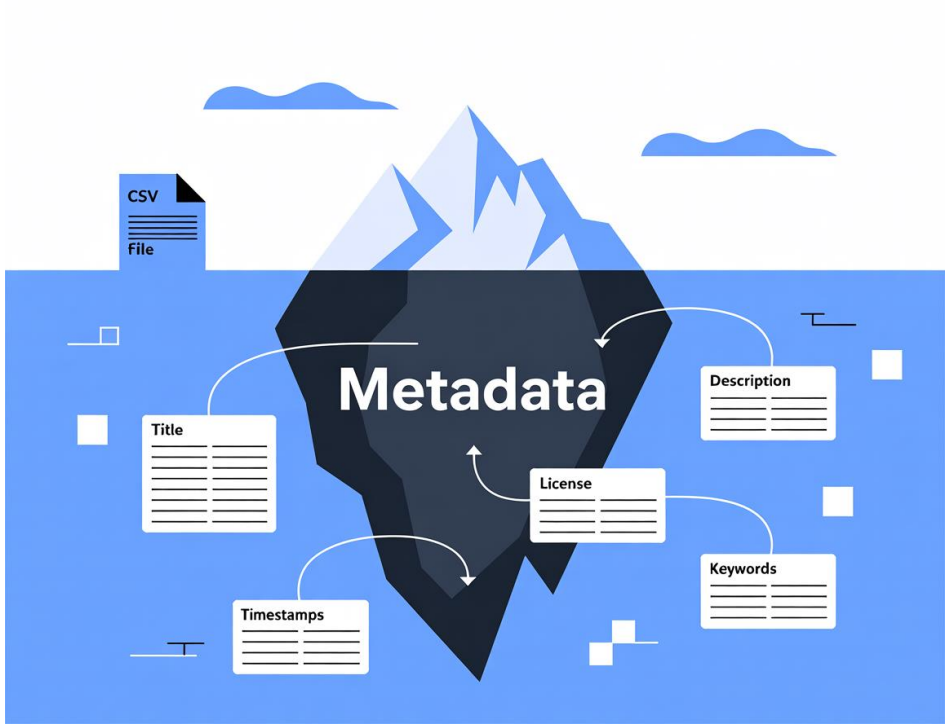
Imagine arriving at the airport baggage carousel after a long international flight. The conveyor belt starts moving and out come hundreds of identical black suitcases. No tags. No labels. No names. No destinations.

At first, people are slightly amused. Then confusion sets in. Passengers start picking up bags, opening them, checking the contents, and putting them back. Someone walks off confidently—only to return a few minutes later, realizing they grabbed the wrong one.

Before long, the entire system slows to a halt. Not because the luggage isn't there—but because no one knows what belongs to whom.

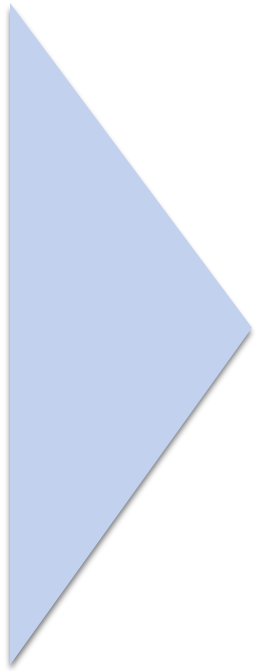
THE METADATA BLIND SPOT

- Enterprises underestimate the importance of metadata, yet: Without it, AI systems lack context and governance becomes impossible (Source: Gartner Report)
- 56% of organizations are investing in data and metadata collection & management (Source: Informatica 2026)
- By 2028, 50% of organizations will adopt "zero-trust" data governance due to unreliable data sources (Source: Gartner 2026)



WHAT IS METADATA (AND WHY IT MATTERS)?

- Provides context, meaning and structure to raw data
- Helps users and systems find, understand, trust and use data



Business - *What is this and why does it matter?*

Technical - *How is this data structured and stored?*

Transactional - *What happened to this data over time?*

Operational- *How is this data used and how well is it working?*

WHAT "AI-READY DATA" REALLY MEANS

Data Quality
Accuracy, Completeness, consistency

Metadata & Context
Business definitions, Data lineage
Semantic layers

Governance
Policies, Ownership, Compliance

Accessibility
Unified, discoverable data

Trust Layer
Observability, Validation, Explainability

Insight

Data becomes a product, not a byproduct

FROM DATA TO INNOVATION: WHAT MUST CHANGE

INVEST IN **DATA QUALITY BEFORE AI SCALE**

TREAT **METADATA AS CRITICAL INFRASTRUCTURE**

ALIGN **AI STRATEGY WITH DATA GOVERNANCE**

THANK YOU