

Failure modes	Root cause	Primary impact
Data retrieval & accessibility gaps	Fragmented sources, restricted access	Incomplete analytics, blind spots in ML features, stale agent decisions
Integration complexity & semantic drift	Mismatched schemas, shifting field definitions	Reports that don't reconcile, models learning wrong patterns
Data quality issues	Duplicates, nulls, inconsistent formats	Unreliable dashboards, unstable ML training, GenAI inheriting defects
Security & compliance constraints	Aggressive PII masking, data residency limits	Compliant but data-starved AI systems
Scalability & performance bottlenecks	Pipelines built for batch, extended to real-time	Failures under volume, broken agentic workflows
Ownership & governance failures	Unclear accountability, undocumented logic	Pipelines that degrade silently over time
Tooling mismatch & technical debt	Legacy Extract, Transform, Load (ETL) extended beyond design scope	Brittle workarounds, undocumented handoffs