

WHITE PAPER

# The Agentic Lakehouse for Financial Services & Insurance

An agentic data platform enables smarter AI, stronger compliance, and dramatically lower costs for banks, capital markets firms, and insurers.

**\$50B**

Global agentic AI market spend in 2025 (KPMG)

**44%**

Of finance teams deploying agentic AI in 2026 (Wolters Kluwer)

**600%+**

Year-over-year growth rate for agentic AI adoption

# The AI Moment Is Here. The Data Foundation Is Not.

*For Chief Data Officers, Heads of Risk Technology, Compliance Leaders, and Data and Analytics executives in the Financial Services and Insurance industries*

Financial services and insurance institutions are entering a defining era. Agentic AI — autonomous systems capable of planning, reasoning, and executing multi-step workflows — is no longer a future concept.

KPMG estimates global agentic AI market spend at \$50 billion in 2025. Wolters Kluwer projects that 44% of finance teams will deploy agentic AI in 2026, a growth rate exceeding 600% year over year. Yet while 99% of financial institutions plan to deploy agents, only 11% have succeeded. The most cited barrier: data.

Fragmented data estates, inconsistent governance, and prohibitive infrastructure costs are blocking the industry from realizing the full value of AI. Legacy data warehouses built for human BI queries cannot deliver the sub-second performance, semantic consistency, and open governance that agentic workloads demand. Dremio changes that.

## The Financial Services Data Crisis

Most financial institutions maintain 10 to 15 disconnected core systems. Customer data lives in one platform, transactions in another, risk models in a third. Integrating this data for fraud detection, Basel III reporting, or real-time AI agents requires brittle ETL pipelines that break under the demands of modern agentic workloads.

**Key stat:** 37% of financial institutions name unstructured, siloed, or poor-quality data as their #1 barrier to achieving AI goals. *(IFC / SME Finance Forum, 2025)*

The cost compounds. Traditional warehouses charge per query and per compute second, creating severe cost overruns when AI agents generate thousands of concurrent requests. Data duplication across warehouses, data marts, and AI silos means the same information is stored and paid for multiple times without improving quality or consistency.

Regulatory pressure amplifies the urgency. Banks operate under Basel III, FRTB, MiFID II, Dodd-Frank, SOX, and DORA. Insurers add IFRS 17, Solvency II, and state privacy mandates. Meeting these obligations with legacy architectures requires armies of data engineers managing brittle pipelines — resources that should be directed at innovation.

# Dremio: The Agentic Lakehouse Built for FSI

Dremio's Agentic Lakehouse is architected from the ground up for AI agents as first-class users. Every layer is designed for the high-concurrency, semantically consistent, governance-enforced workloads that FSI agentic AI demands.

## One Governed View — Without ETL

Query live transaction, KYC, and market data in a single view with no ETL and no data movement. Data stays in low-cost cloud object storage (S3, ADLS, GCS), eliminating the expensive copy-and-move patterns that define legacy architectures. A fraud agent, a risk reporting agent, and a Customer 360 dashboard can all query the same governed data simultaneously.

## AI Semantic Layer

A living, governed encyclopedia of the business. Define metrics once — "Suspicious Transaction," "Customer Risk Score," "Regulatory Capital Ratio," "Claims Frequency" — and every AI agent, dashboard, and regulatory report consumes identical definitions. When regulators examine a Suspicious Activity Report, the risk score driving it matches the one used by compliance, model risk, and internal audit. Consistency is automatic and auditable.

## Autonomous Reflections

As AI agents generate thousands of concurrent queries, Dremio's Active Metadata engine autonomously reorganizes data layout and creates intelligent pre-aggregations — no DBA tuning required. Result: 10 to 100x faster query performance at any workload volume.

## AI Agent & MCP Server

A built-in MCP Server lets any external AI model (ChatGPT, Claude, Gemini, or a proprietary enterprise model) connect, query the AI Semantic Layer, and execute governed workflows with full audit trails. No custom integrations. No governance gaps. One open standard, any model.

## Enterprise Security & Compliance

SOC 2 Type 2, ISO 27001, HIPAA-ready. Fine-grained column and row-level security, dynamic data masking, complete data lineage, customer-managed encryption keys, and data residency controls for GDPR and DORA — all native, not add-ons.

## Open Architecture: No Lock-In

Data is stored in Apache Iceberg format in your own cloud account. Any compatible engine (Spark, Trino, Flink) queries the catalog directly via Apache Polaris and the Iceberg REST Catalog spec. As AI models evolve, your data platform evolves with them.

Dremio delivers immediate, measurable value across the highest-priority AI use cases in financial services and insurance.

Use Case	How Dremio Helps	Business Outcome
<b>Real-Time Fraud &amp; AML</b>	Millisecond queries on live and historical data; unified AI semantic layer; ML feature serving without ETL	Faster interdiction, fewer false positives, lower fraud losses
<b>Risk Aggregation &amp; Regulatory Reporting</b> (Basel III, FRTB, Solvency II)	Virtualized golden sources; intraday VaR/LCR/RWA; AI agent-automated report generation	65% reduction in reporting cycle time; audit-ready outputs
<b>Insurance: Claims, Underwriting &amp; Pricing</b>	Unified policy, claims, actuarial, telematics, and IoT data; dynamic pricing models	More accurate pricing, faster claims resolution, lower combined ratio
<b>Customer 360 &amp; KYC Onboarding</b>	Real-time federated customer profiles; governed PII handling; AI-assisted document extraction	30% to 50% faster onboarding; reduced operational cost
<b>Portfolio &amp; Investment Analytics</b>	Sub-second analytics on live market and position data; open format for quant models	Faster alpha generation, better risk-adjusted decisions

*Customer references: FactSet (20x faster financial data delivery), Nomura (faster, more stable risk analytics), Vanguard (AI semantic layer supporting \$8T AUM), TD Securities (governed data fabric across 75 silos).*

# The Dremio Advantage

For FSI institutions, the data platform decision is one of the most consequential technology choices of this decade. The platform serving as the foundation for agentic AI must be purpose-built for it — not adapted from a batch BI warehouse.

Capability	Dremio Agentic Lakehouse	Legacy Warehouse / Cloud DW
<b>Data format</b>	Open Apache Iceberg with no lock-in; data stays in your cloud account	Proprietary or semi-open; vendor controls data
<b>AI/Agent integration</b>	Built-in MCP server, native AI agent, AI Semantic Layer	Bolt-on add-ons; no AI semantic layer context for agents
<b>Query federation</b>	Native with no ETL, no data movement across all sources	Manual ETL pipelines required; data duplication
<b>Cost structure</b>	\$0.20/DCU, elastic, no per-query fees; 70 to 90% lower costs	Per query / per credit; unpredictable at AI scale
<b>Performance tuning</b>	Autonomous Reflections: self-optimizing; 10 to 100x faster	Manual indexing; dedicated DBA effort required
<b>Compliance readiness</b>	SOC 2 Type 2, ISO 27001, HIPAA-ready; full lineage and audit trails	Varies; often requires additional tooling
<b>Deployment speed</b>	Minutes to provision; approximately 90 days to full production ROI	Months to years; complex migrations required

# Business Case & Path to Value

Dremio's implementation model is designed to generate value quickly without requiring a "big bang" migration. Most enterprise customers achieve full production deployment within 90 days and realize ROI within their first quarter.

**Days 1–30: Connect and Activate** Connect to 2 to 3 high-priority data sources, define initial semantic objects for the most critical use case (fraud scoring or a specific regulatory report), and deploy the AI Agent for analyst self-service. Zero data migration required.

**Days 31–90: Extend and Automate** Extend to additional source systems, deploy first autonomous agent workflows (AML monitoring, intraday risk aggregation, or automated compliance reports), and activate Autonomous Reflections as usage patterns emerge.

**Day 91+: Enterprise Scale** Enterprise-wide data access, multi-agent orchestration for complex end-to-end workflows, full cost benefits as legacy ETL pipelines are decommissioned. Most customers achieve measurable ROI within this phase.

Business Impact Area	Typical Result
Infrastructure cost reduction	70% to 90% lower vs. legacy warehouse
Query performance improvement	10 to 100x faster via Autonomous Reflections
Regulatory reporting cycle time	65%+ reduction through AI agent automation
Client onboarding / KYC speed	30% to 50% faster through automated workflows
Time to production deployment	Approximately 90 days; ROI typically within first quarter
Agentic AI ROI	2.3x average return within 13 months (KPMG, 2025)

**Open Standard Advantage:** As a co-creator and leading contributor to Apache Arrow, Apache Iceberg, and Apache Polaris, Dremio sits at the center of the open data ecosystem. Data stored in open Apache Iceberg tables in your own cloud account remains portable across engines and AI tools — your insurance policy against vendor lock-in and platform obsolescence.

Financial services and insurance institutions need a data platform built for the agentic era — not adapted from yesterday's BI infrastructure. Dremio delivers four capabilities that no legacy platform can match simultaneously: unified data access without ETL, governed AI semantics enforced at the platform layer, autonomous self-optimizing performance, and open architecture with zero vendor lock-in.

Dremio is the only FSI-validated platform with native MCP integration, open Iceberg storage, and zero-ETL federation. It is the lowest-risk choice for institutions serious about governance and the most cost-efficient choice for institutions serious about their bottom line. The window for competitive differentiation through agentic AI is open — but not indefinitely.

## About Dremio

Dremio is the Agentic Lakehouse company. Dremio Cloud gives data teams the ability to run fast, self-service SQL analytics on open data lake storage, eliminating data warehouse lock-in and ETL complexity. Trusted by leading financial services institutions, asset managers, insurance companies, and global banks, Dremio is built on Apache Iceberg and Apache Polaris to ensure data portability and interoperability across the entire modern AI ecosystem.

## Sources

1. KPMG Global Agentic AI Report, 2025
2. Wolters Kluwer Finance Teams & Agentic AI Study, 2026
3. IFC / SME Finance Forum AI Barriers Study, 2025
4. Forrester / AWS Agentic AI Readiness Survey, 2025
5. Neurons Lab, "Agentic AI in Financial Services: A Research Roundup for 2026"

# Ready to accelerate your AI journey?

Dremio is the only FSI-validated platform with native MCP integration, open Iceberg storage, and zero-ETL federation. See it in action.

[Request a Demo](#)