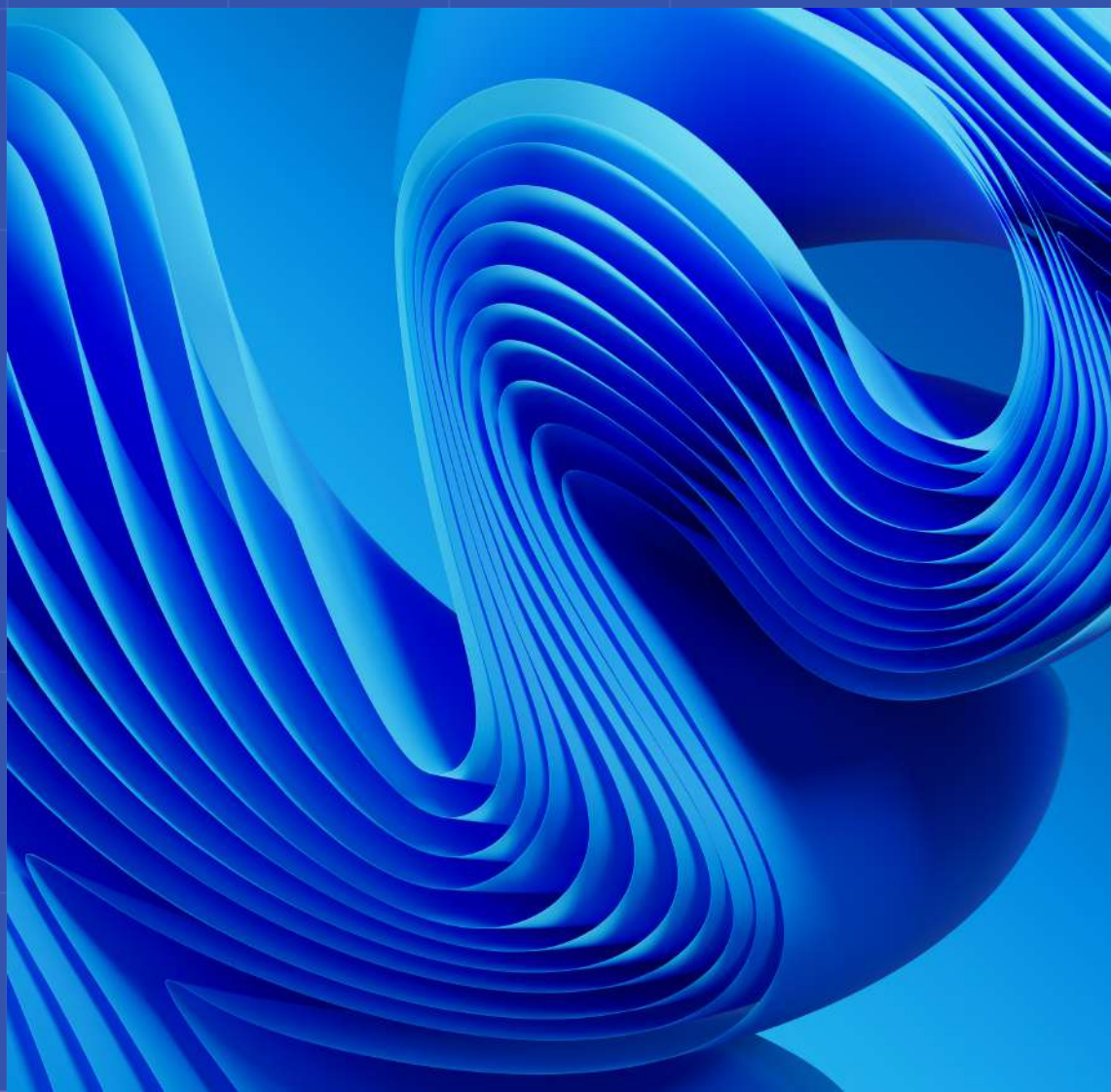
DataArt

From Legacy Burden to AI Advantage

A 2025 Roadmap for Capital Markets Data Modernization

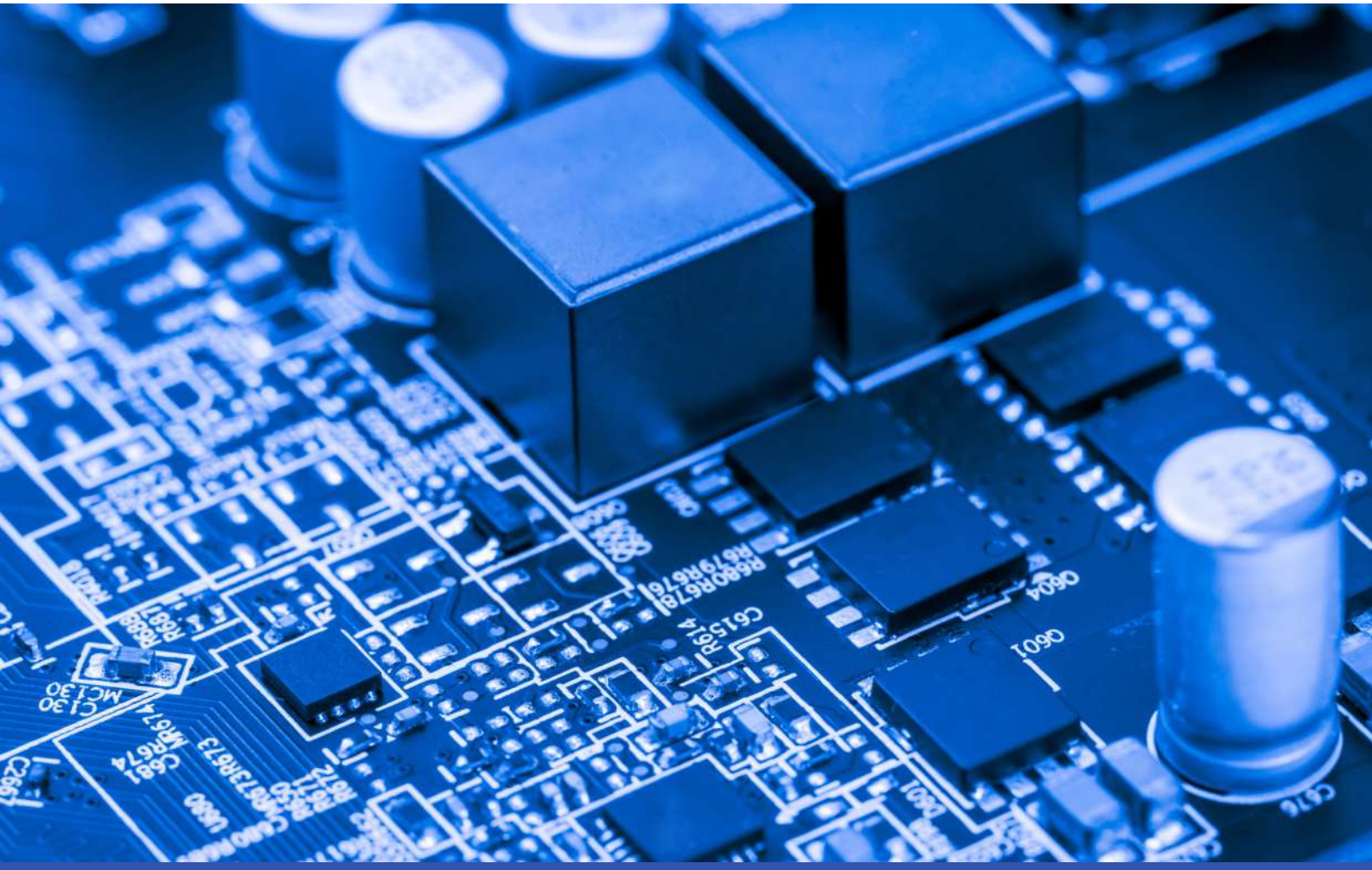


Executive Summary

Capital markets firms stand at an inflection point where legacy trading systems—built as isolated silos over decades—are constraining both operational efficiency and AI innovation. The real breakthrough isn't just moving to the cloud or implementing AI models; it's architecting modern data platforms that can **incrementally replace legacy components** while **stitching together existing systems** to unlock immediate value.

This whitepaper explores how modern data architectures enable a **piece-by-piece transformation** strategy, leveraging **federated domain models** with simplified interfaces and **agentic AI** that serves dual purposes: accelerating the software development lifecycle (SDLC) and delivering direct business value through autonomous workflows.

DataArt specializes in this holistic approach—from serverless environments and lake architectures to agentic SDLC processes and business-value creation—enabling firms to modernize without the risks of "big bang" replacements.



1 — The Trading Systems Reality: System of Record vs. System of Intelligence

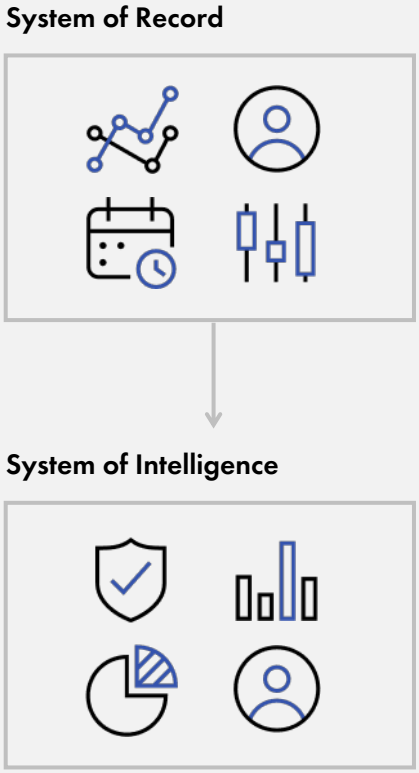
Most capital markets firms operate with a fundamental architectural split that legacy approaches struggle to bridge:

System of Record: Real-Time Trading Core

- **Trade execution engines** handling microsecond-latency requirements
- **Position management systems** maintaining real-time book state
- **Risk engines** providing instant limit checks and portfolio monitoring
- **Market data feeds** processing millions of ticks per second

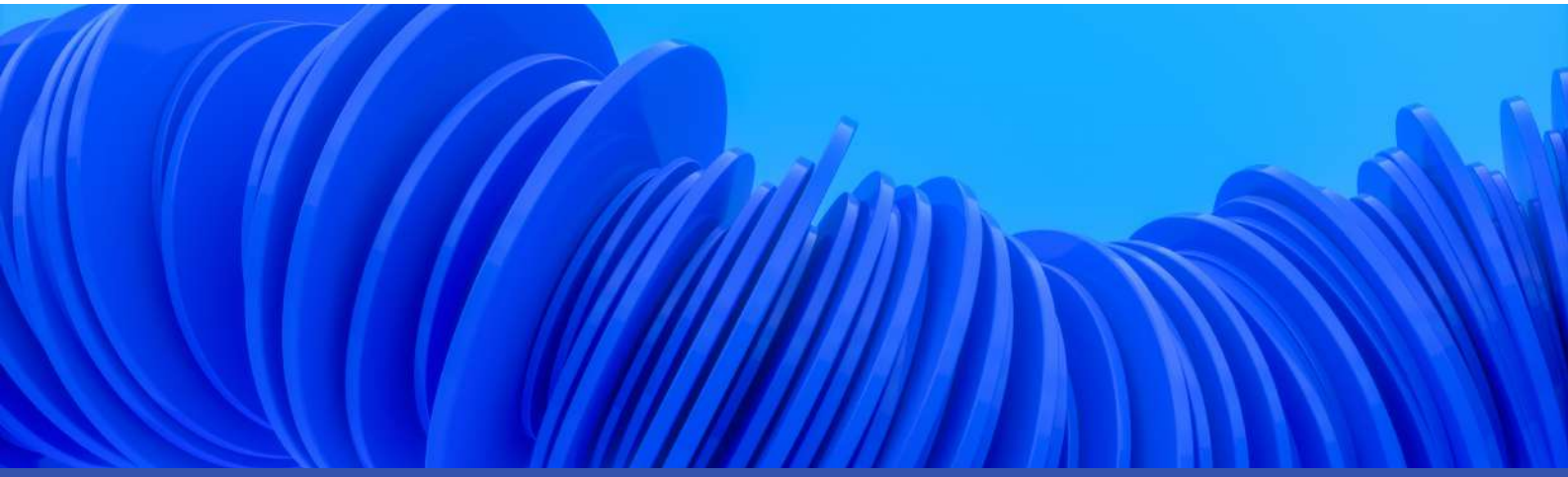
System of Intelligence: Compliance & Analytics Scaffold

- **Regulatory reporting** systems requiring full data lineage
- **Risk aggregation** platforms for Basel III, stress testing
- **Surveillance systems** monitoring for market abuse
- **Client analytics** providing insights and recommendations



The Challenge: Legacy architectures force a choice between real-time performance and analytical depth. Trading systems optimize for speed but lack the rich context needed for compliance and AI. Analytics systems have the data breadth but can't operate at trading speed.

The Modern Solution: A unified data fabric that maintains ultra-low latency for trading operations while simultaneously feeding rich, contextualized data streams to compliance, risk, and AI systems—without requiring changes to core trading engines.



2 — Incremental Modernization: The Strangler Fig Strategy

The transformative power of modern data architecture lies not in wholesale replacement but selective enhancement and intelligent integration.

Replace One Piece at a Time

Rather than ripping out entire trading platforms, modern data architectures enable:

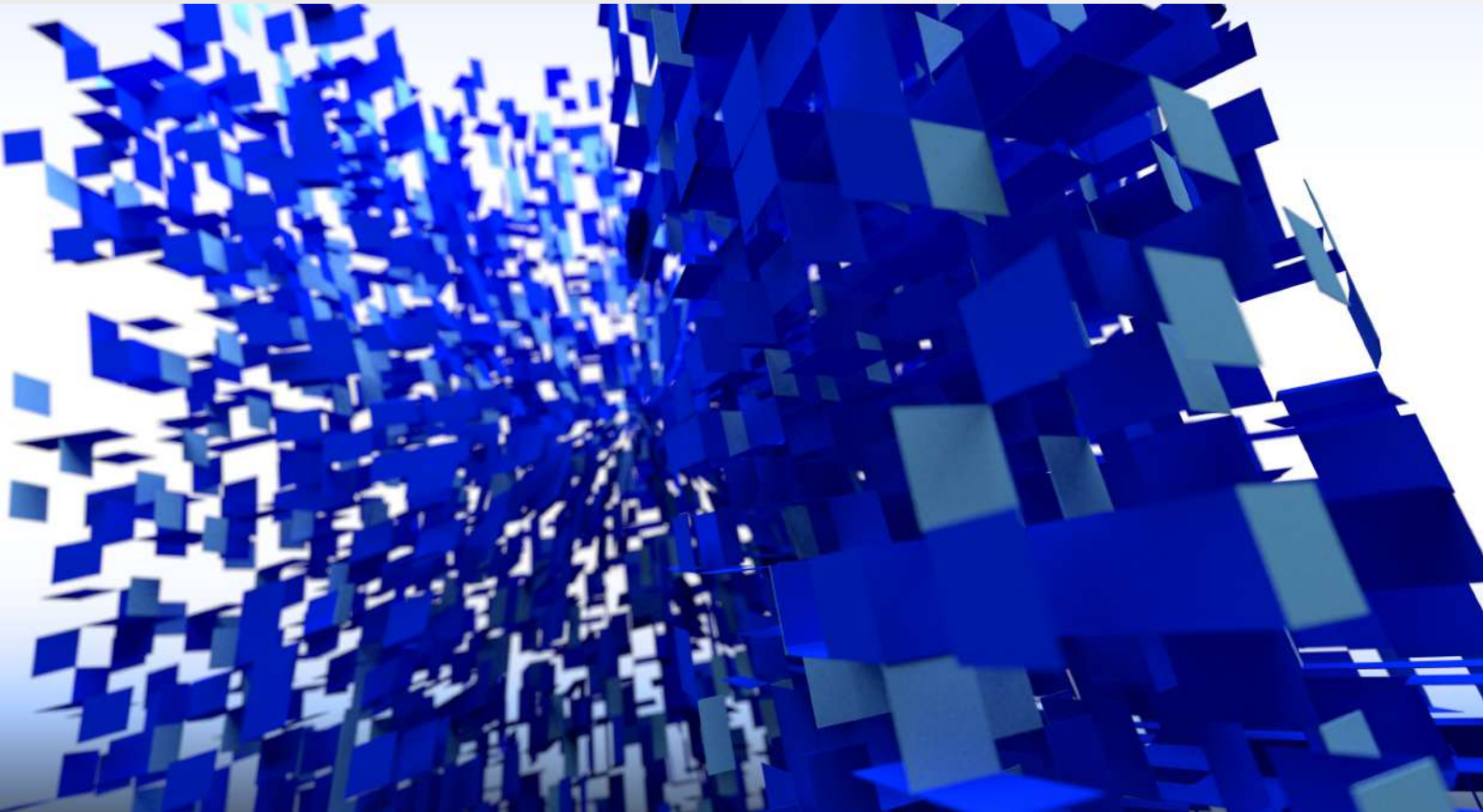
- **Wrap legacy position systems** with real-time streaming APIs while keeping core logic intact
- **Replace aging surveillance systems** with AI-powered alternatives that consume the same data feeds
- **Modernize reporting engines** piece by piece, maintaining backward compatibility
- **Upgrade risk systems** incrementally, validating against legacy calculations during transition

Stitch Systems Together for Immediate Value

Modern platforms excel at creating **cross-system intelligence** without disrupting existing operations:

- **Combine trade data with client communications** for enhanced surveillance without touching either core system
- **Merge position data with market research** to create personalized client insights
- **Integrate settlement data with credit systems** for real-time exposure monitoring
- **Connect trading patterns with compliance alerts** for proactive risk management

This approach delivers immediate ROI while building toward a more integrated future.



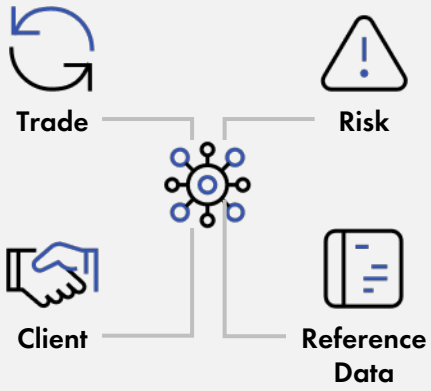
3 — Federated Domain Models: Simplifying Complex Integrations

Traditional integration approaches require deep knowledge of each system's internal data model and complex ETL processes. **Federated domain models** revolutionize this by establishing:

Standardized Business Interfaces

Instead of integrating at the technical level, modern architectures define **business-domain interfaces:**

- **Trade Domain:** Standard trade lifecycle events (new, modify, cancel, fill)
- **Risk Domain:** Common risk measures and limit definitions
- **Client Domain:** Unified client profiles and interaction history
- **Reference Data Domain:** Instrument definitions and corporate actions



Event-Driven Integration

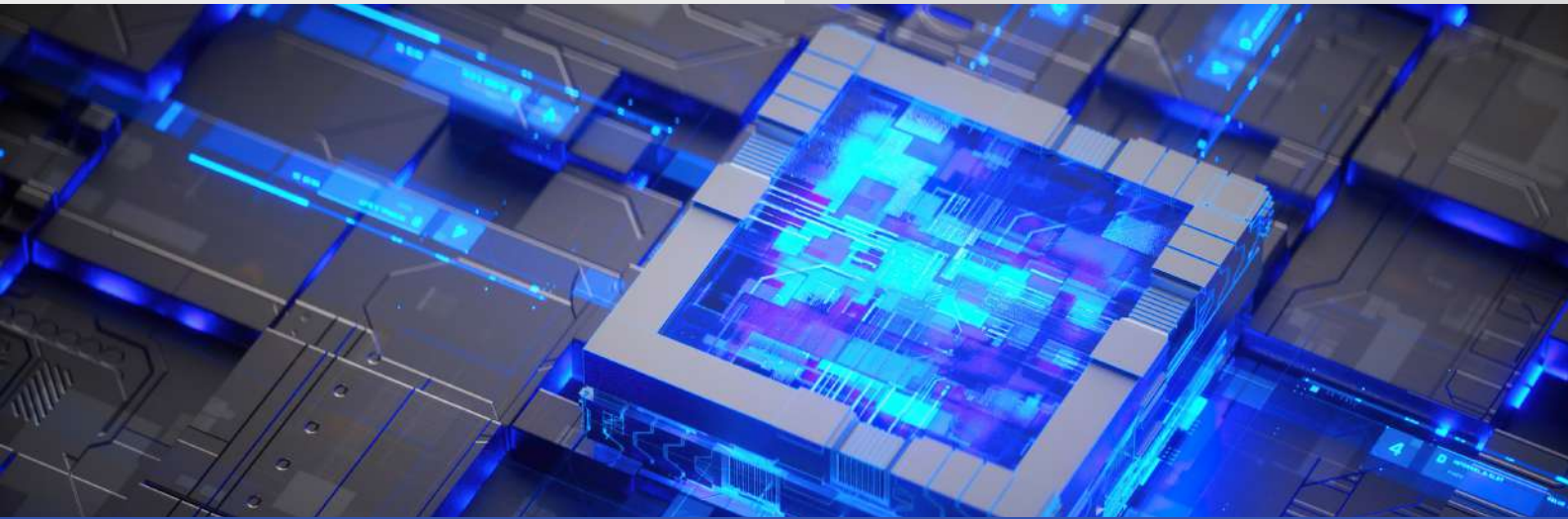
Each domain publishes **business events** in standardized formats, enabling:

- **Loose coupling** between systems—changes to one don't break others
- **Easy interface development**—new systems can consume standard events
- **Real-time synchronization** across all connected systems
- **Audit trails** automatically maintained through event logs

API-First Design

Modern interfaces are **much easier to write** because they:

- Use **REST/GraphQL APIs** instead of custom protocols
- Provide **self-documenting schemas** with automatic validation
- Support **versioning** for backward compatibility
- Enable **microservice integration** without system-wide dependencies



4 — Agentic AI: Dual-Purpose Transformation

Agentic AI represents a paradigm shift beyond traditional machine learning models. These autonomous agents can **reason, plan, and execute tasks** with minimal human intervention, serving two critical functions:

Agentic SDLC: Accelerating Development

AI-Powered Development Agents transform how data platforms are built and maintained:

- **Code Generation Agents:** Automatically create data pipelines, API endpoints, and integration layers based on business requirements
- **Testing Agents:** Generate comprehensive test suites, perform data quality validation, and execute regression testing
- **Documentation Agents:** Maintain real-time documentation, API specs, and architectural diagrams
- **Monitoring Agents:** Proactively identify performance issues, data quality problems, and security vulnerabilities

Impact: Development cycles that previously took months can be compressed to weeks, with higher quality and fewer bugs.

Agentic Business Value: Direct Operational Impact

- **Business Process Agents** operate autonomously within the data fabric:
- **Trade Surveillance Agents:** Continuously monitor trading patterns, automatically flagging suspicious activities and adapting detection rules based on new market behaviors
- **Client Engagement Agents:** Analyze client portfolios and market conditions to proactively suggest trade ideas, rebalancing opportunities, or risk adjustments
- **Compliance Agents:** Monitor regulatory changes, automatically update reporting rules, and ensure all trades meet current requirements
- **Settlement Agents:** Predict and resolve potential settlement failures before they occur

Key Advantage: These agents operate in **real-time** across the federated domain model, accessing data from multiple systems while respecting business rules and compliance requirements.

5 — Modern Data Platform Architecture: The Technical Foundation

The platform enabling this transformation combines several architectural principles:

Serverless-First Design

- **Event-driven compute** that scales automatically with market activity
- **Cost optimization** through pay-per-use pricing models
- **Zero infrastructure management** overhead for development teams
- **Instant scalability** during market volatility or end-of-day processing

Lake Architecture with Domain Segmentation

- **Raw data layer** maintaining immutable audit trails
- **Domain-specific data marts** optimized for particular business functions
- **Real-time streaming** for latency-sensitive applications
- **Batch processing** for complex analytics and regulatory reporting

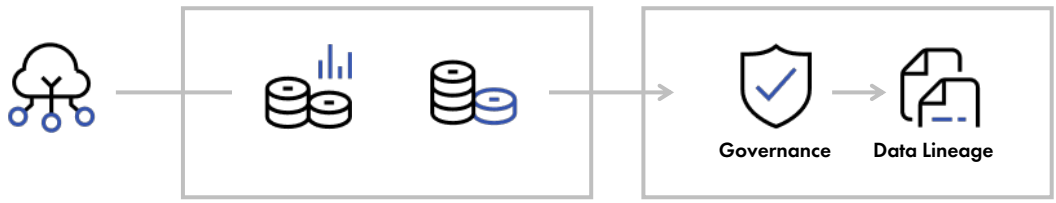
Embedded Governance and Lineage

- **Automatic data cataloging** with business context
- **End-to-end lineage tracking** for regulatory compliance
- **Policy-based access controls** respecting business domain boundaries
- **Data quality monitoring** with automatic remediation

Multi-Cloud Resilience

- **Cross-cloud data replication** for operational resilience
- **Vendor-agnostic architectures** avoiding lock-in
- **Regional data residency** compliance for global operations
- **Disaster recovery** with sub-minute recovery times

Modern Data Platform Architecture



6 — Phased Implementation Framework

Phase 1: Foundation and Quick Wins

(Months 1-3)

1. **Establish a serverless data lake** for consolidated logging and basic analytics
2. **Implement streaming infrastructure** to capture real-time events
3. **Deploy the first agentic SDLC tools** for automated testing and documentation
4. **Identify high-value integration opportunities** requiring minimal system changes

Phase 2: Domain Federation

(Months 4-6)

1. **Define federated domain models** for core business areas
2. **Implement API gateways** with standardized interfaces
3. **Deploy first business-value agents** (e.g., enhanced surveillance)
4. **Begin incremental replacement** of selected legacy components

Phase 3: Advanced Integration

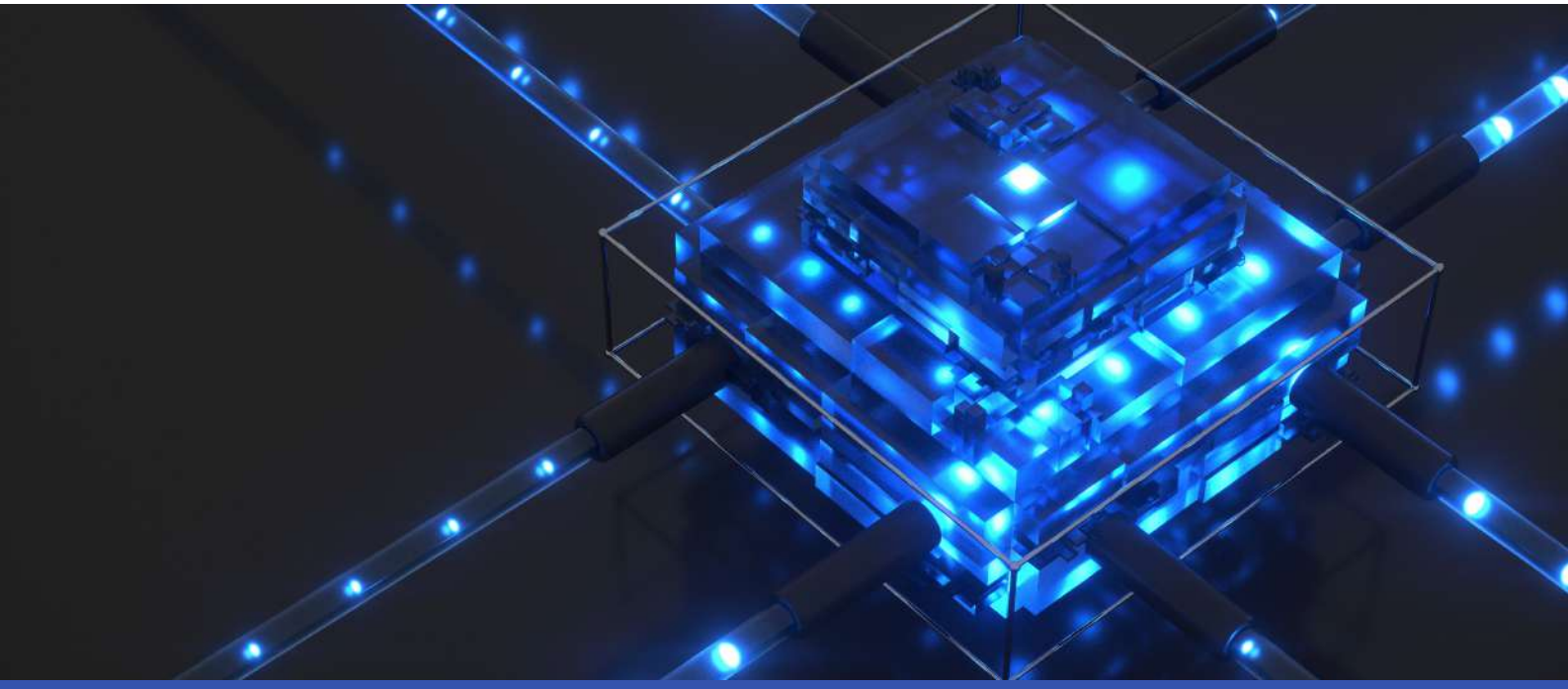
(Months 7-12)

1. **Expand cross-system intelligence** capabilities
2. **Deploy sophisticated agentic workflows** for client engagement and compliance
3. **Implement full regulatory reporting** through a modern stack
4. **Establish a center of excellence** for ongoing expansion

Phase 4: Optimization and Scale

(Months 12+)

1. **Fine-tune agentic AI** based on operational learnings
2. **Expand to additional domains** and business units
3. **Implement advanced AI capabilities** (predictive analytics, autonomous trading)
4. **Prepare for next-generation use cases** (digital assets, DeFi integration)



7 — DataArt's Comprehensive Approach

DataArt brings unique capabilities spanning the entire transformation spectrum:

Serverless Infrastructure Expertise

- **Cloud-native architectures** optimized for financial services
- **Serverless computing patterns**, reducing operational overhead
- **Event-driven designs**, supporting real-time requirements
- **Multi-cloud strategies** for resilience and flexibility

Lake Architecture Implementation

- **Data lakehouse designs** combining flexibility with performance
- **Streaming-first pipelines** for real-time data processing
- **Domain-specific optimization** for trading, risk, and compliance workloads
- **Governance frameworks** meeting regulatory requirements

Agentic SDLC Integration

- **AI-powered development tools**, accelerating delivery
- **Automated testing frameworks**, ensuring quality
- **Continuous integration/deployment** pipelines for data platforms
- **DevOps best practices** adapted for financial services

Business-Value Agentic AI

- **Financial domain expertise**, ensuring agents understand business context
- **Regulatory-compliant AI**, meeting audit and explainability requirements
- **Real-time decision engines**, operating at trading speeds
- **Human-AI collaboration** frameworks for optimal outcomes

Capital Markets Domain Knowledge

- **Deep understanding** of trading systems and market structure
- **Regulatory expertise** across global jurisdictions
- **Risk management** specialization for various asset classes
- **Industry relationships** enabling best-practice sharing

8 — Regulatory Alignment and Future-Proofing

Basel III Endgame Readiness

Modern architectures naturally support:

- **Full data lineage** for risk data aggregation
- **Real-time capital calculations** enabling dynamic optimization
- **Comprehensive audit trails** for regulatory examination
- **Stress testing capabilities** with scenario modeling

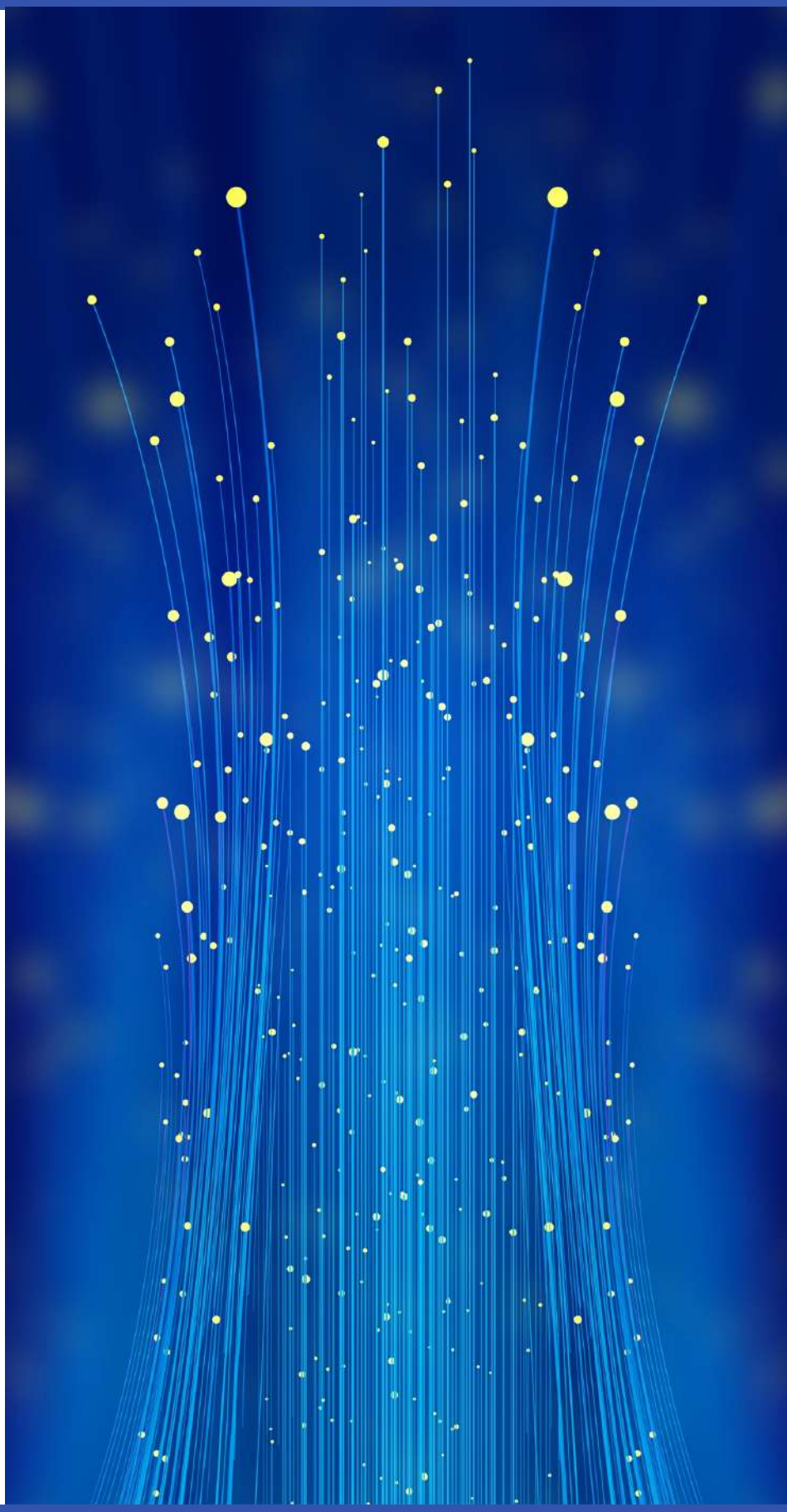
AI Governance Framework

Built-in support for:

- **Model explainability** through data lineage tracking
- **Bias detection and mitigation** across trading algorithms
- **Performance monitoring** of AI-driven decisions
- **Regulatory reporting** on AI system usage and outcomes

Operational Resilience

- **Multi-region failover** capabilities
- **Real-time backup and recovery** for critical systems
- **Cyber threat detection** through behavioral analytics
- **Business continuity** planning with automated responses



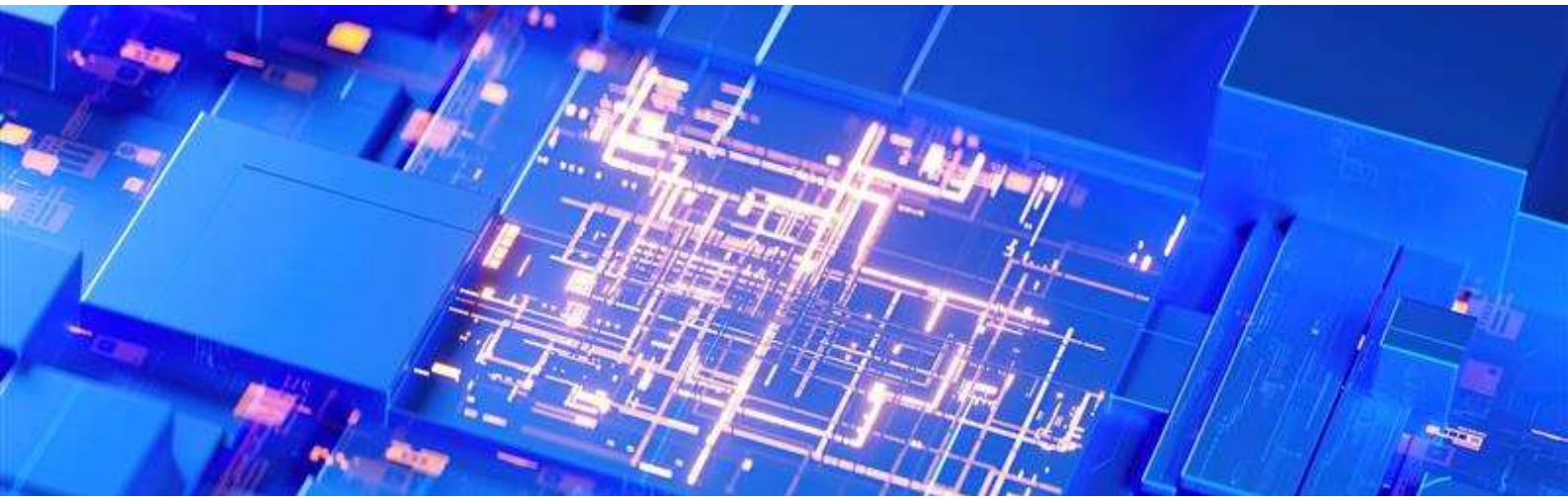
Conclusion: The Strategic Imperative

The capital markets industry is experiencing a fundamental shift. Firms operating with fragmented legacy systems will find themselves increasingly disadvantaged as AI-native competitors enter the market and regulatory expectations evolve.

The path forward is clear: **incremental modernization** through modern data architectures enabling **piece-by-piece transformation** while **immediately unlocking cross-system value**. The combination of federated domain models and agentic AI provides both the technical foundation and business acceleration needed to compete effectively.

The **moment to begin is now**. Market leaders are already implementing these approaches, and the competitive gap will only widen as agentic AI capabilities mature and regulatory requirements intensify.

DataArt stands ready to partner with forward-thinking capital markets firms in this transformation, bringing proven expertise across infrastructure modernization, agentic AI implementation, and regulatory compliance—enabling the journey from legacy burden to AI advantage.



Contact DataArt to explore how federated domain models and agentic AI can transform your capital markets operations while respecting your existing system investments.

Learn more about DataArt's capital markets solutions at **www.dataart.com/insights**