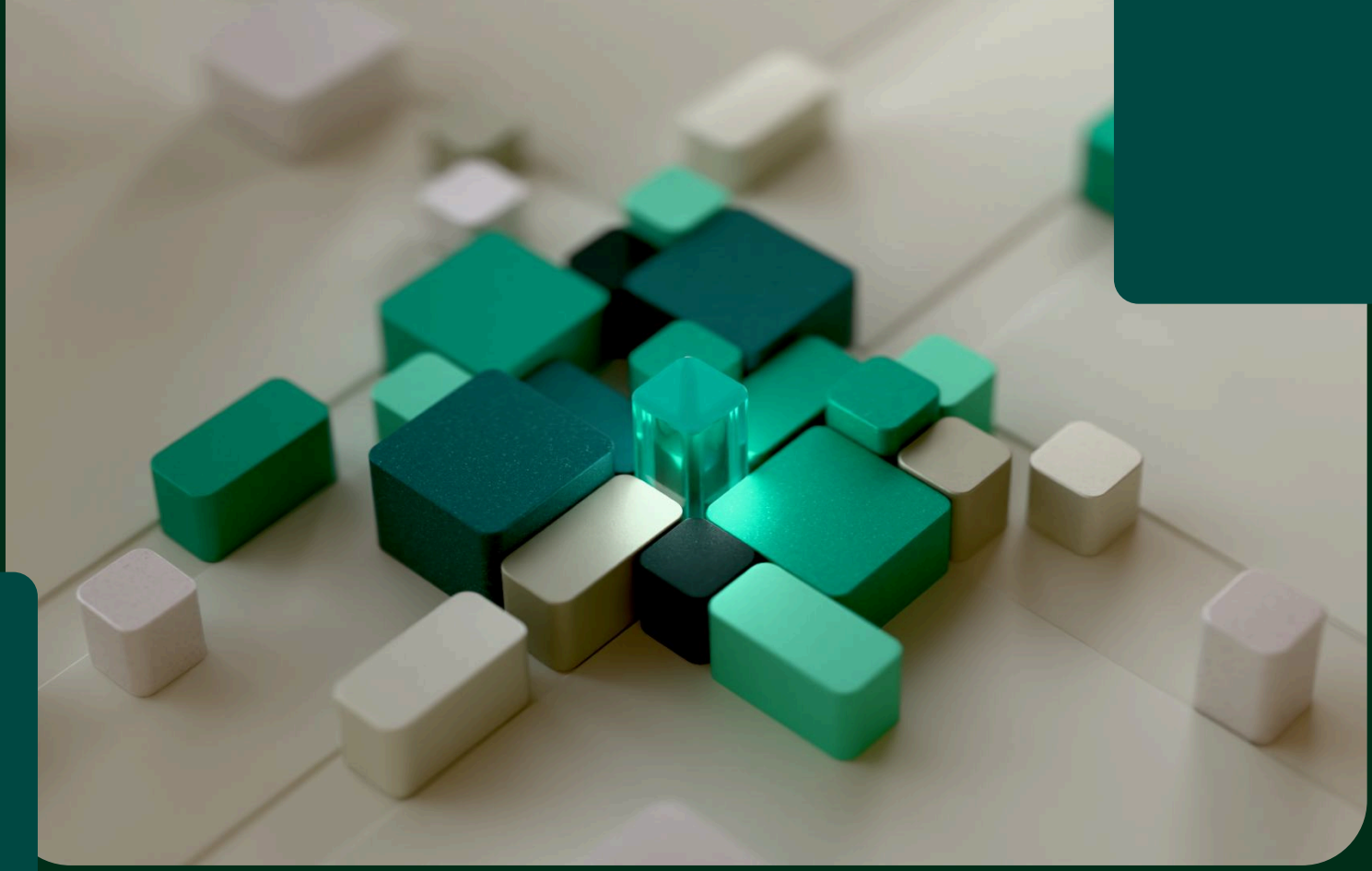

DataHub

Building BCBS 239 Compliance:

Strengthen your data foundation for risk and AI

A GUIDE TO BUILDING AUDIT-READY RISK DATA INFRASTRUCTURE FOR MODERN FINANCIAL INSTITUTIONS

DataHub

How DataHub enables scalable risk data governance and AI-ready compliance

Context gives meaning to data

More than a decade after its introduction, BCBS 239 remains one of the most critical frameworks shaping how financial institutions manage risk data. Yet, as of late 2023, only two banks were deemed fully compliant with all 14 principles.[1] Legacy IT systems, fragmented governance, inconsistent data quality, and underinvestment continue to hinder progress. Non- compliance carries real consequences—from regulatory sanctions and capital restrictions to reputational damage and operational disruption.

But compliance is not the endgame. When approached strategically, BCBS 239 offers a rare opportunity: the same controls that satisfy regulators also build a foundation for faster decision-making, operational efficiency, and trustworthy AI adoption.

This guide shows how DataHub helps financial institutions move beyond checking the compliance box to building resilient, future-ready data foundations. By mapping BCBS 239 principles directly to DataHub's metadata-native capabilities, we demonstrate how banks can reduce compliance risk today while accelerating transformation for tomorrow.

What you'll learn in this guide

This guide takes a principle-by-principle approach, showing how DataHub can help banks not only meet each of the requirements of BCBS 239, but use it as the basis of driving data-preparedness for AI, improving CX, and achieving overall business agility.

Specifically, we'll explore how DataHub helps financial institutions ensure they stay ready for BCBS 239 audits with the strategy and technology to support them.

BCBS 239 compliance is more than a regulatory mandate—it's a catalyst for transformation. With DataHub, financial institutions can not only meet supervisory expectations but also unlock operational efficiency, resilience, and the trustworthy data foundation required for AI at scale.

DataHub

Meeting BCBS 239 with data context:

A principle-by-principle guide

1. Overarching governance and infrastructure

Related BCBS-239 principles:

Principle 1: Governance

BCBS 239 requires banks to establish a strong governance framework for risk data aggregation and reporting. Boards and senior management must take clear ownership of risk data processes, ensuring defined responsibilities, effective oversight, and alignment with enterprise risk management.

Principle 2: Data architecture and IT infrastructure

Banks must maintain an IT and data architecture that supports accurate, consistent, and scalable risk data aggregation.

This means designing systems that allow databases to communicate seamlessly and ensuring they can capture, aggregate, and report risk data effectively. It also calls for establishing shared knowledge graph across the organization—defining metadata, applying consistent identifiers, and standardizing naming conventions for entities such as legal organizations, customers, and accounts.

The challenge

BCBS 239 begins with a strong foundation: governance and infrastructure. Yet for many banks, these remain the hardest gaps to close.

- **Siloed ownership** across lines of business and functions creates inconsistent controls and unclear accountability.
- **Legacy systems**—often fragmented and not designed for unified aggregation—make it difficult to establish a single view of risk.
- **Inconsistent terminology** across domains leads to confusion in regulatory reporting, where precision and alignment are critical.

DataHub

> "The overall pace of banks' progress in implementing sustainable risk data aggregation and risk reporting capabilities is occurring at a slower pace than envisaged. This is largely because several banks have persistent challenges with fragmented IT landscapes, legacy systems and manual processes that are not fit for purpose."
>
> - Progress in adopting BCBS 239, November 2023

How DataHub helps with governance and infrastructure

A unified, coherent view of risk data

DataHub is built on a knowledge graph foundation, acting as a **single pane of glass** that ingests metadata across all banking platforms—databases, warehouses, streaming systems, BI tools and proprietary applications.

- Custom connector framework integrates with legacy and specialized systems.
- "Catalog of catalogs" ingests and harmonizes metadata from existing catalogs, avoiding redundant migrations.
- Federated metadata service provides a real-time, enterprise-wide view.

Why it matters: Banks gain a coherent, always-current understanding of all assets relevant to risk.

Governance workflows and data contracts

DataHub's built-in workflows ensure assets meet governance standards before they are shared.

- Role-based annotations and structured approval processes document compliance.
- Metadata quality standards ensure ongoing completeness and consistency by guiding how people participate in data management. They provide checks and workflows that hold individuals accountable for documenting, reviewing, and approving assets.
- Data contracts formalize sharing agreements and provide an audit trail of approvals.

Why it matters: Banks can demonstrate to regulators that governance is embedded in every step of the data lifecycle, with clear evidence of who approved what, when, and under which conditions.

Consistent data definitions

DataHub's **centralized business glossary** capabilities support standardization, while maintaining flexibility.

- Institutions can define critical terms like "exposure," "customer," or "default" in a consistent, controlled way.
- Glossaries are federated by domain and subdomain, reflecting how the bank actually operates (e.g., Retail Banking vs. Corporate Banking) without forcing a rigid, one-size-fits-all standard.

Why it matters: This structure eliminates ambiguity in risk reporting while allowing each domain to retain the business-specific definitions it needs.

Improved data ownership and accountability

DataHub organizes assets by **domains** and publishes them as **data products.**

- Ownership is assigned within each business domain (e.g., a "Credit Risk Data Product" owned by the credit risk function), fostering accountability.

Why it matters: Every dataset has a clear owner and steward, creating accountability for quality and compliance.

The bottom line

BCBS 239 starts with governance and infrastructure because they set the foundation for everything else. With DataHub, banks can meet these principles by embedding governance into daily operations, assigning clear ownership, unifying metadata visibility, and standardizing risk definitions—without forcing a rigid, one-size-fits-all glossary.

2. Risk data aggregation capabilities

Related BCBS-239 principles:

Principle 3: Accuracy and integrity

BCBS 239 requires risk data to be accurate, reliable, and free from errors. Banks must define data consistently, monitor accuracy, and implement escalation plans for any discrepancies.

Principle 4: Completeness

Banks must be able to capture and aggregate all relevant risk data across the organization to build a holistic view of risk exposures and emerging threats.

Principle 5: Timeliness

Risk data aggregation must be reported promptly to support decision-making. Decision makers must have access to up-to-date information, especially under stress conditions.

Principle 6: Adaptability

Risk data processes must be flexible enough to handle new risks, regulatory expectations, and ad hoc requirements, particularly during times of market volatility or stress.

> "Several banks still lack a common taxonomy and complete data lineage, which further complicates banks' ability to harmonize systems and detect data defects."
>
> - Progress in adopting BCBS 239, November 2023

The challenge

While principles 3–6 are clear, banks often struggle to operationalize them at scale.

- **Incomplete or inconsistent** data across business lines undermines accuracy.
- **Manual validation workflows** obscure data origins and increase error risk.
- **Rigid batch pipelines** delay reporting, especially under stress conditions.
- **Static architectures** make it difficult to adapt quickly to new regulatory or reporting requirements.

How DataHub helps with risk data aggregation

Proactive governance and quality assurance

DataHub embeds quality controls directly into the metadata layer, enabling proactive governance and early
detection of issues.

- Shift-left governance with automated tests for data quality rules, schema consistency, and compliance with data standards.
- Issues are detected directly in the platform before they reach production.

Why it matters: Banks prevent inaccurate data from entering risk systems, improving audit readiness and strengthening confidence in risk calculations.

End-to-end data lineage for comprehensive aggregation

Lineage in DataHub gives banks a complete, auditable
view of how risk data flows across systems and transformations.

- Cross-platform, column-level lineage automatically traces data flows from source systems through transformations, aggregations, and calculations and into final risk reports.
- Business context, definitions, and data quality rules are consistently applied and understood across the entire data supply chain.

Why it matters: Banks ensure that all risk data is captured, complete, and traceable back to authoritative sources, ensuring no critical information is overlooked.

Streamlined compliance workflows

DataHub replaces manual documentation and approvals with automated, auditable workflows.

- Automated documentation and approval processes for data access, changes, and usage.
- Generates auditable records that regulators expect
- Compliance forms let governance teams crowdsource documentation and classification, ensuring consistent metadata quality and faster compliance readiness.

Why it matters: Banks reduce manual compliance burdens and provide regulators with clear, auditable evidence of governance.

Controlled access to sensitive data

DataHub enforces strict access policies with full transparency.

- Granular access controls limit exposure of sensitive risk data to only authorized users, systems, and even AI models.
- All activity is logged with full audit trails.

Why it matters: Banks protect sensitive risk data and demonstrate compliance with data privacy and security requirements.

How DataHub helps with risk data aggregation

Real-time metadata for operational agility

With event-driven ingestion, DataHub creates a real-time view of data assets to support rapid reporting and stress scenarios.

- Low-latency ingestion from Kafka streams and event sources keeps metadata current.
- Metadata change proposal system designed for low-latency updates ensures rapid updates.

Why it matters: Banks support real-time risk assessment and reduce delays in risk reporting.

Scalable metadata infrastructure

Designed for global enterprises and financial institutions, DataHub can process metadata at scale without slowing down reporting.

- High-volume metadata ingestion designed for global organizations handling massive datasets ensures the system doesn't become a bottleneck for timely reporting.

Why it matters: Banks ensure metadata is always available and up to date, even in high-volume reporting environments.

Flexible data models for evolving requirements

DataHub's extensible metadata model makes it easy to capture new dimensions and regulatory attributes.

- Extensible metadata models allow banks to capture new risk dimensions, classifications, and rules without disrupting existing systems.

Why it matters: Banks adapt quickly to new regulatory requirements and ad hoc reporting needs. This is especially critical during periods of crises or market volatility.

Programmatic access for dynamic reporting

DataHub's robust APIs and SDKs give banks direct, flexible ways to integrate metadata into custom tools and reporting.

- APIs and SDKs allow teams to programmatically query and interact with metadata.
- Teams can build custom risk reporting applications and integrate with existing systems to generate dynamic reports.

Why it matters: Banks can respond rapidly to new supervisory requests and integrate with advanced analytics platforms for deeper risk insights.

The bottom line

Principles 3–6 demand accuracy, completeness, timeliness, and adaptability in risk data aggregation. With DataHub, banks can embed these requirements directly into their metadata fabric—delivering high-quality, governed data at scale and in real time.

The result: resilient compliance today and agility to meet the regulatory and business challenges of tomorrow.

3. Risk reporting practices

Related BCBS-239 principles:

Principle 7: Accuracy

BCBS 239 requires that risk reports reflect an accurate, clear, and truthful state of the bank's risk profile. Banks must be able to provide an inventory of the validation rules applied to quantitative information (i.e. mathematical or logical relationships).

Principle 8: Comprehensiveness

Risk reports should capture all material risk exposures, including emerging risks, with recommendations for action. They must allow decision makers to understand the full scope of exposures and position information for all significant risk areas (i.e. credit risk, market risk, liquidity risk, operational risk).

Principle 9: Clarity and usefulness

Reports should be clear, concise, and tailored to the audience. They must balance detail with interpretation and provide actionable insights for management and boards.

Principle 10: Frequency

Reports must be produced at intervals that match the bank's risk profile and external conditions. In times of stress, more frequent reporting may be required to give stakeholders near-real-time visibility.

Principle 11: Distribution

Risk information must be delivered to the right stakeholders at the right time, ensuring those responsible for oversight have what they need to act.

> "Supervisors expect that in times of stress/crisis all relevant and critical credit, market and liquidity position/exposure reports are available within a very short period of time to react effectively to evolving risks."
>
> - BCBS 239,

The challenge

Meeting principles 7–11 requires banks to deliver accurate, timely, and actionable reports at scale. But many face persistent roadblocks:

- **Inaccurate or outdated reports** undermine trust in decision making.
- **Lack of semantic alignment** leads to inconsistencies across business units.
- **Inflexible delivery mechanisms** hinder reporting frequency and responsiveness.
- **Opaque data flows** make it difficult to trace reported metrics back to their source.

How DataHub helps risk reporting practices

Continuous data quality assurance

DataHub embeds accuracy checks directly into the metadata layer to ensure quality at the source.

- Continuous metadata tests and "shift-left" governance validate data quality rules, schema consistency, and governance standards directly within the platform, ensuring data is accurate at its source and throughout its lifecycle.
- Automated data quality checks catch issues early in the reporting lifecycle.
- Data health dashboards provide visibility into data quality coverage, incidents, and trends, enabling teams to triage issues quickly, monitor failures over time, and ensure consistent reliability across critical datasets.

Why it matters: Banks reduce the risk of inaccurate data being used in risk calculations and reporting.

Consistent data definitions

DataHub ensures reporting is based on a shared understanding of risk concepts across the enterprise.

- Centralized business glossary ensures consistent terminology and definitions for risk-related data, reducing ambiguity and potential for errors in interpretation.
- Glossaries can be federated by domain and subdomain while still aligning to enterprise-wide standards.

Why it matters: Banks eliminate ambiguity in reports. They ensure metrics are understood and interpreted consistently across business units.

Auditable data lineage

DataHub gives banks an end-to-end, auditable trail from raw source data through transformations into final reports.

- Cross-platform, column-level lineage provides a clear and auditable trail of data from source to report.
- Data transformations and calculations are fully documented and verifiable.

Why it matters: Banks validate the accuracy of reports, providing regulators with clear evidence of compliance.

The bottom line

Principles 7–11 focus on the reporting layer: accuracy, comprehensiveness, clarity, frequency, and distribution. With DataHub, banks can embed these principles into their reporting fabric—ensuring that risk reports are accurate, consistent, and timely, with lineage that proves their reliability. Banks deliver reports stakeholders trust, even under regulatory or market pressure.

4. Supervisory review, tools, and cooperation

Related BCBS-239 principles:

Principle 12: Review

Supervisors are expected to conduct ongoing evaluations of a bank's compliance with BCBS 239. They may also issue ad hoc requests on specific risk topics with tight turnaround times, as a way to test the bank's ability to quickly aggregate risk data and generate accurate reports.

Principle 13: Remedial actions

If banks fall short on risk data practices, supervisors are expected to step in with corrective measures. These may include requiring remediation, intensifying supervision, or mandating independent reviews. In more serious cases, supervisors can impose capital add-ons as both a safeguard and an incentive for compliance.

Principle 14: Cooperation

Banks that operate across multiple jurisdictions require coordination between home and host supervisors. This cross-border cooperation enables effective global oversight and fosters information sharing among regulators. The result is more consistent and aligned enforcement of BCBS 239 standards.

> "At certain banks, board and senior management lack awareness/attention to data issues, and therefore do not ensure appropriate budget, resources and accountability for risk data aggregation and reporting initiatives."
>
> - Progress in adopting BCBS 239, November 2023

The challenge

Principles 12–14 emphasize transparency and accountability. Yet banks often lack the infrastructure to deliver them consistently:

- **Missing audit trails** obscure governance actions and ownership.
- **Gaps remain unresolved** because there is no systematic way to track remedial steps.
- **Fragmented documentation** makes it difficult for supervisors to gain timely access to the evidence they need.

How DataHub helps with supervisory review, tools, and cooperation

Controlled access and monitoring

DataHub enforces strong access governance and creates transparency around every interaction with sensitive data

- Granular access controls ensure only authorized users can modify or access sensitive risk data.
- Audit trails record every action for full visibility.

Why it matters: Banks safeguard risk data from tampering and provide supervisors with clear evidence of data integrity.

End-to-end lineage for verification

Lineage in DataHub gives supervisors and auditors confidence that data has not been altered improperly.

- Cross-platform, column-level lineage tracks the complete flow of data from source to report to ensure it has not been compromised during its journey.
- Every transformation and calculation is fully documented.

Why it matters: Banks confirm the integrity of reported data and enable independent verification by regulators.

Scalable infrastructure for frequent reporting

DataHub is designed to handle large-scale metadata ingestion and updates without compromising timeliness.

- High-volume ingestion supports rapid capture of metadata across large, complex estates.
- Low-latency updates ensure metadata stays current even during stress events.

Why it matters: Banks generate risk reports as frequently as required without infrastructure bottlenecks. Supervisors gain confidence that reporting processes can scale under pressure.

The bottom line

Principles 12–14 focus on supervisory oversight, remedial action, and cooperation. With DataHub, banks operationalize these expectations through robust audit trails, end-to-end lineage, and scalable reporting infrastructure. Banks demonstrate transparency, accelerate remediation, and build regulator trust by making compliance evidence accessible and verifiable.

DataHub

Beyond Compliance

Turning BCBS 239 into a strategic advantage

Compliance that pays dividends—if you do it right

Too often, BCBS 239 is seen as a box to check. But with the right metadata foundation, it can unlock durable, compounding benefits for your organization. Implementing the governance, quality, and transparency requirements of BCBS 239 doesn't just satisfy regulators—it lays the groundwork for:

- **Operational efficiency** through automation
- **Faster time-to-insight** via self-serve discovery
- **AI/ML readiness** with trustworthy, explainable data
- **Future compliance agility** with extensible metadata models

In other words: the same controls that mitigate risk also create opportunity.

Governance → automation

Establishing domain ownership and data contracts through DataHub turns governance into an operational accelerator.

- **Automated data contract enforcement** ensures policies are applied consistently.
- **Shift-left governance at scale** embeds controls early in the data lifecycle.
- **Federated data product management** empowers domains to own their data while aligning with enterprise standards.

Banks reduce manual governance overhead and scale risk data processes with greater efficiency.

Data quality + lineage → AI-readiness

When data meets BCBS 239 standards for accuracy, completeness, and traceability, it becomes a foundation for AI.

- **Reliable for AI models** that depend on trustworthy inputs.
- **Auditable for regulators and internal review,** ensuring AI outputs can be explained.
- **Accessible and extendable via APIs and the DataHub Model Context Protocol (MCP) Server,** enabling metadata-aware AI agents.

Banks build confidence that AI models are not only powerful, but also transparent, compliant, and safe to deploy.

Standardization → future-proofing

By modeling risk definitions, glossary terms, and data flows in DataHub, banks set themselves up for agility.

- **Eliminate ambiguity** across domains with consistent, federated terminology.
- **Accelerate adoption of new regulations** by reusing established metadata structures.
- **Handle ad hoc stress scenarios** without rework, thanks to extensible metadata models.

Banks gain the ability to adapt faster than the market and stay ahead of regulatory demands.

Compliance Requirement	Competitive Edge
Lineage — Trace data from source to report	Explainability — AI/ML outputs can be trusted and audited
Domain ownership — Assign accountable stewards	Automation — Scalable enforcement of governance policies
Standardization — Consistent terms, identifiers, and taxonomies	Agility — Faster adoption of new regulations and stress scenarios
Data contracts — Formal agreements for data usage	Operational efficiency — Reduced manual review and approvals
Quality controls — Continuous metadata tests	AI-readiness — Reliable, complete inputs for advanced models
Audit trails — Document every approval and change	Transparency — Regulator and stakeholder trust

The bottom line

BCBS 239 was designed to close the gaps exposed during the financial crisis, but its principles remain just as critical today. Too often, compliance has been treated as a box-ticking exercise—a costly and reactive effort to satisfy regulators.

But the reality is this: **the same capabilities that ensure compliance are the ones that future-proof the business.**

With DataHub, banks gain both the strategy and the technology to prepare for BCBS 239 reporting. Governance workflows, end-to-end lineage, quality checks, and federated glossaries are no longer abstract policies—they are operationalized in a metadata-native platform built for scale. That means:

- **Reporting is audit-ready by design,** with traceable lineage and documented approvals.
- **Supervisory reviews are accelerated,** with clear evidence of governance at every stage.
- **Risk data aggregation and reporting are resilient,** even under market stress or evolving regulations.

Looking ahead, the banks that thrive will be those that turn regulatory requirements into launchpads for transformation—adopting modern AI and data context platforms to power not just risk management, but also greater innovation, advanced analytics, and superior customer experience.

By embedding governance, trust, and context at the core, you prepare your organization not just for today's compliance challenges, but for tomorrow's competitive opportunities.

Take the next step toward confident compliance and innovation

Start building your risk data advantage with DataHub.
Share a few details about your goals and current challenges, and we'll schedule a personalized walkthrough tailored to your specific data environment, use cases, and roadmap.

Contact us: sales@datahub.com

Compliance doesn't have to be a cost center.
With DataHub, it becomes a strategic advantage.

DataHub

About DataHub

DataHub, by Acryl Data, is an AI & Data Context Platform. Innovated jointly with a thriving open source community of 13,000+ members, DataHub's active metadata platform provides real-time context of AI and data assets with best-in-class scalability and extensibility. The company's enterprise SaaS offering, DataHub Cloud, delivers a fully-managed solution with AI-powered discovery, observability, and governance capabilities. Organizations rely on DataHub to accelerate time-to-value from their data investments, ensure AI system reliability, and implement unified governance—enabling AI & data to work together and bring order to data chaos.

Learn more at datahub.com
Follow DataHub on LinkedIn and X