

WHITE PAPER

BCBS 239 Compliance & Graph Technology

How Data Connections Enable More
Cost-Effective Financial Risk Reporting

Nav Mathur, Senior Director, Global Solutions

As governmental regulations tighten, today's banks must have a thorough and systematic understanding of risk calculations and their associated data lineage—including where underlying data originates and how it flows through enterprise systems.

This paper discusses how connected data and graph database technologies are transforming risk reporting in modern banks to help them meet the stringent demands of risk reporting compliance.

At the same time, those forward-looking banks are uniting data silos into an information foundation for building innovative applications. These solutions provide extreme visibility and deep analytical insights that improve compliance efforts and day-to-day decision making.

White Paper

TABLE OF CONTENTS

The Connected Nature of Financial Risk	1
The Emergence of Risk Reporting Standards	1
BCBS 239 Regulation Principles	2
Key Challenges of BCBS 239 Regulations	3
Using a Federated Approach to BCBS 239 Compliance	4
Building a Risk Metadata Foundation	4
Choosing the Right Graph Technology	5
The Finance Industry Business Ontology	6
Real-World Risk Reporting Solutions	8
BCBS 239 and Graph Technology Are Game Changers	9
A Foundation for Innovative Applications	9

BCBS 239 Compliance & Graph Technology

How Data Connections Enable More Cost-Effective Financial Risk Reporting

Nav Mathur, Senior Director, Global Solutions

The Connected Nature of Financial Risk

The lack of timely risk data was a major contributor to the global financial crisis of 2008 as the collapse of Lehman Brothers sent shockwaves through the banking world. Without standards for properly aggregating risk in their financial positions, banks were unable to quickly assess the dependency of their various holdings on Lehman stock and assets.

Armed with an understanding of risk *data lineage* – a visibility of data connections all the way back to authoritative data sources – financial houses could have limited their exposure. Such visibility requires financial data standards and modern software that understands the connectedness of modern investment instruments.



The Emergence of Risk Reporting Standards

Since the 2008 market meltdown, [regulators have established standards](#) for recording and tracing financial transactions and for aggregating risk data. The new standards are designed to uncover risk dependencies and adjust capital ratio requirements accordingly.

To create consistency in recording financial contract details, the International Organization for Standardization (ISO) released ISO 17442, the Legal Entity Identifier (LEI) initiative. LEI codes clearly identify parties in transactions, thereby laying a sturdy foundation for deep visibility into financial risk data.

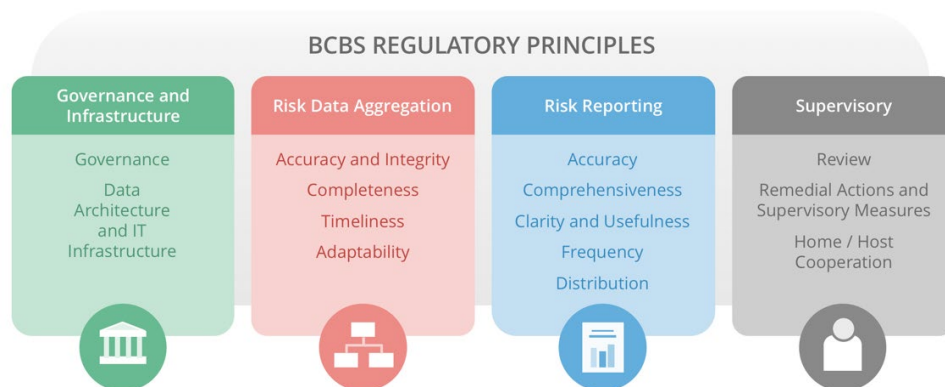
Another crucial data-focused initiative is BCBS 239, the Basel Committee's fourteen principles to be used by banks when aggregating financial risk data.

These new standards collectively enable banks to assess risk, [trace data lineage](#), and understand dependencies on other systems, investments and financial houses.

BCBS 239 regulations require banks to build a risk management infrastructure based on industry standards and repeatable, auditable processes.

BCBS 239 Regulatory Principles

The fourteen BCBS 239 principles are organized into four data management categories.



Governance and Infrastructure

To build their risk reporting systems, banks must utilize data governance and integrated data taxonomies, as well as group-wide metadata including consistent identifiers for entities, counterparties, customers and accounts. The banks must also maintain data systems that handle the requirements of normal operations as well as the high demands and specific requirements of crisis situations.

Risk Data Aggregation

Banks must be able to generate accurate, consistent and reliable risk data while maintaining full visibility back to authoritative data sources. Any aggregations and transformations must adjust for data latency, so all calculations are based on data values from the same point in time. And the datasets must be able to satisfy a full spectrum of requests made by managers and regulators.

Risk Reporting

Management and regulatory reports must represent risk in a precise and auditable manner, and reconcile to the complexity of the bank's risk model and operations. They must present risk information in a clear, concise and easily understood manner than facilitates fast, informed decisions. The reports must be distributed regularly to managers and regulators and also be available for on-demand and ad hoc requests.

Supervisory Review

Bank risk supervisors are required to review their institution's ongoing compliance with BCBS 239 principles. They must have access to the tools required to address any deficiencies they discover in their investigations. Supervisors are also required to cooperate with other regulators and other supervisors in their compliance investigations and implementation of remedial actions.

BCBS 239 requirements create complex data management challenges that require new, innovative approaches and technology.

Key Data Challenges of BCBS 239 Regulations

Building models for risk reporting requires tackling some serious data management issues.

Data Lineage

Accurate and reliable risk reports require a clear understanding of [data lineage](#). Reporting entities must be able to prove how each number in a report is generated, including its calculation details and source data. Each data item and transformation must be attributed to an owner, a steward, and be profiled with a quality and latency status. Most importantly, data must be traced backwards until its lineage ends with an authoritative source.

Data Silos

Lines of business often grow their own systems to meet specific business needs or to create independent trading desks for faster decision making. This produces discrete [data silos](#) that make tracing data lineage a very difficult task. You must be able to trace data movement through those silos and systems all the way back to their original sources. While data warehouses can assemble information from discrete silos, they do little or nothing to trace data lineage, and can even make the process more complex.

Terminology Differences

Business groups often use their own terminology and algorithms, even within the same organization. For example, the *notional value* of a derivative contract can mean different things to different people. What is a derivative? Exactly what asset classes are included? Is the data original, copied or calculated? Is the data derived from internal or external sources? Are those sources authoritative?

Legal Entity Identifiers

With the introduction of the Legal Entity Identifier Act (LEI) and MiFID 2,¹ entities and counterparties in contracts are required to use standard identifiers to describe transactions. While this helps address accountability of the parties, developers must still relate old entity identifiers to the new identifiers on all historical information.

Data Consistency and Latency

The data appearing in regulatory reports must be accurate and consistent as of a specific time. To achieve such temporal consistency, all transactions must be timestamped. Reports must use those timestamps to assemble accurate snapshots of risk data at any point in time.

Achieving consistency across data silos presents even harder challenges. Risk reports typically pull data from multiple sources, each with its own refresh schedule. To avoid data latency issues and achieve consistency, risk reports must adjust for temporal differences in each silo.

1 The Markets in Financial Instruments Directive (MiFID) is the framework of European Union (EU) legislation for: investment intermediaries that provide services to clients around shares, bonds, units in collective investment schemes and derivatives (collectively known as "financial instruments").

Complexity issues are leading banks to employ federated models that leave data in its original location and use graph databases and centralized metadata to address data lineage requirements.

Using a Federated Approach to BCBS 239 Compliance

Integrating information into a single, enterprise-wide logical data model is very difficult and time consuming. In some cases, the structure and location of much of the data makes it all but impossible to address in a single, centralized data store. And ironically, moving everything into a single repository can make tracing data lineage even more difficult.

After earlier failed attempts to centralize enterprise information in data warehouses and operational data stores, most banks have accepted that their data will remain in silos.

Given these complications, many banks are now embracing a federated model that leaves the data dispersed in its original locations, while maintaining control of the model using centralized metadata.



[Federated metadata models](#) make it considerably easier to relate entity identities, maintain data consistency, and describe end-to-end data lineage.

Building a Risk Metadata Foundation

Banking institutions have no choice but to address BCBS 239 regulations. But rather than reactively responding to the new mandates, progressive banks are using BCBS 239 as justification for building a strong metadata foundation for risk management, regulatory and analytic applications.

Traditional metadata management technologies might appear to be an obvious choice for building your new metadata foundation. But they are not capable of handling highly-connected risk-management data, tracing data lineage, or adjusting for temporal inconsistencies in reports.

The complexity of the risk management and BCBS 239 require more than just simple metadata managers. To tackle the modeling and management requirements of the new regulations, [you need to use graph database technology](#).

Neo4j is the leading graph database and is far superior to traditional database technologies for financial data lineage applications.

Choosing the Right Graph Database Technology

Data management experts agree that metadata challenges should be solved using graph databases, and not old, traditional technologies. But all graph technologies are not the same; many are thin veneers built atop old relational or [NoSQL engines](#) with inherent problems.

Relational Databases Aren’t Graph Ready

Relational databases masquerading as graph technologies are fraught with systemic faults. As complex queries traverse a graph model, query hops translate into a flurry of relational table joins that use computing resources inefficiently and cripple application performance. In sharp contrast, native graph databases use graph methods to store and query graph-based metadata. The result is fast, consistent data retrieval, presentation and management.

Non-Native NoSQL Databases Provide No Solution

Non-native NoSQL databases also fall short of core BCBS 239 requirements. Instead of storing data relationships as native graph elements, they add a graph translation layer that reduces query performance. And NoSQL engines lose transactions and relationships regularly, making them unreliable for tracing data lineage back to original information sources.

Neo4j: Native Graph Platform Ready for Risk Compliance

The most popular and successful graph database is [Neo4j](#), which is used in a large majority of graph installations worldwide. As a 100% native graph database, Neo4j eliminates the data consistency and corruption problems caused by non-native approaches to graph applications. And its dependable query performance delivers instant results, even for Value at Risk, Potential Future Exposure and other complex risk-reporting requests.

A Modern Graph Approach is Far Superior for Risk Management

Compliance and Risk Management Tasks	Traditional Approaches	Modern Neo4j Approach
Trace data through enterprise systems	Complex queries with hundreds of join tables	Simple single query traverses all enterprise systems
Preserve the integrity of data lineage	Broken data paths and lineage, especially with NoSQL databases	Continuous, unbroken data paths at all times
Effort required to add new data and systems	Days to weeks to rewrite schema and queries	Minutes to draw new data connections
Time to deployment	Months to years	Weeks to months
Response time for complex queries	Minutes to hours per query	Milliseconds per query
Form of responses	Text reports that are not visual and prove very little	Data lineage visuals show the path it follows through systems
Bottom line	Long, ineffective and expensive	Easy, fast and affordable

By choosing Neo4j for risk reporting compliance, you get a lot more than the world's leading graph database. Neo4j supports global financial terminology standards is backed by professional services that guarantee success and provide new visibility into your compliance efforts and day-to-day operations.

The FIBO ontology is crucial for describing risk data lineage, and graph modeling is perfect for designing and communicating financial risk models.

The Finance Industry Business Ontology (FIBO)

FIBO is a collaborative effort to define and evolve a set of standard terms for investment instruments, business entities, market data, legal obligations and corporate actions affecting global financial markets.

Authored by The Enterprise Data Management Council (EDM), the FIBO ontology has given rise to a series of technical standards governed by the Object Management Group (OMG).

The Benefits of FIBO Terminology

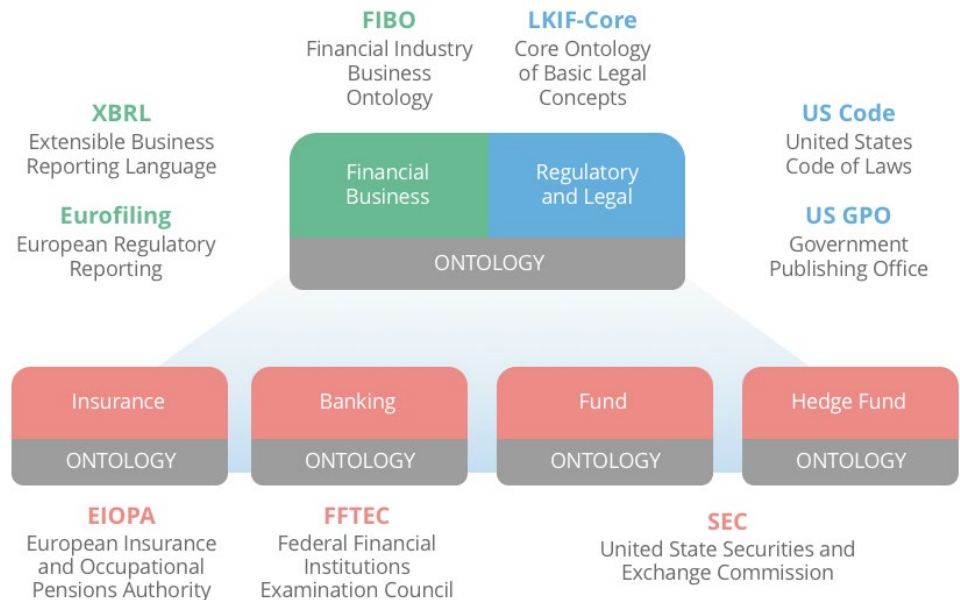
FIBO's great strength stems from its ability to clearly and completely describe the entities, instruments and relationships involved in financial transactions. Such clarity enables financial organizations worldwide to:

- Align data elements across multiple data repositories and silos to achieve data consistency
- Trace the lineage of investment data back to its original, authoritative sources
- Use a standard, common language for communications between financial houses, and between business and technical audiences
- Build rigorous, robust solutions for financial reporting and compliance

FIBO Is Best Represented in Graphs

FIBO is a conceptual model that is best represented as a graph ontology, as depicted by the visual below.

Example: How FIBO Terms Appear in Graph Models



To simplify reporting and compliance requirements, the financial sector is evolving a common language of business and regulatory terminology.

Keeping Pace with Industry Change

The flexibility of FIBO and graph technology makes them excellent standards for the ever-changing financial industry. Their extensibility allows them to adapt as financial markets, technology, industry players and regulations evolve. As the FIBO standard spreads across the industry, the pressure on banks to express risk metadata as a graph of standard language is mounting rapidly.

Neo4j is FIBO-Ready

The FIBO ontology is now available as add-on to the Neo4j platform, so you can create an enterprise canonical data model that:

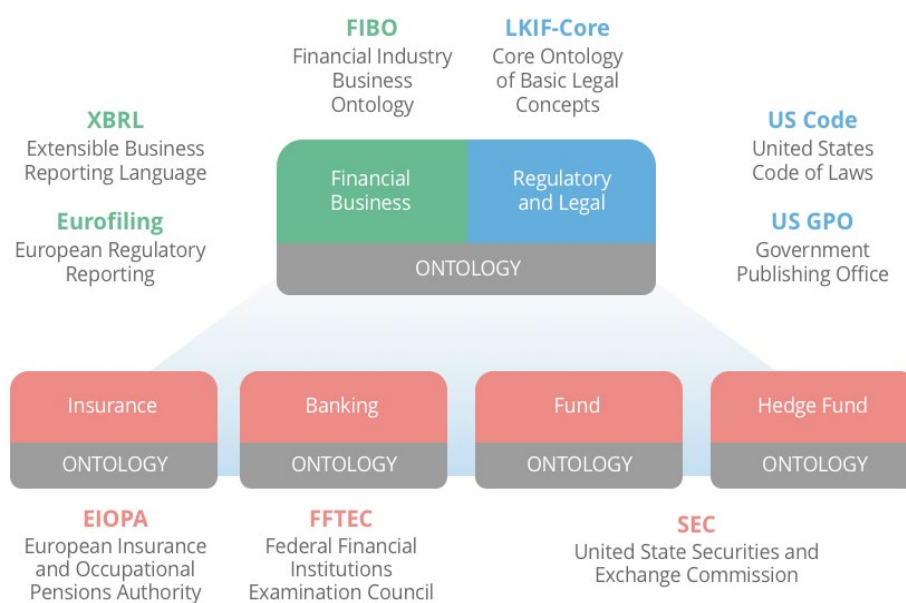
- Uses the same infrastructure to store risk data lineage as well as governance metadata such as definitions, related terms, etc.
- Provides a unified access layer that spans data silos

As a result, by using Neo4j, you can integrate data governance, compliance reporting and real-time data movement into a single solution that guarantees data consistency across operational and regulatory systems.

Handling All Aspects of Compliance

Bank compliance with risk regulations spans business, financial and legal domains. The Legal Knowledge Interchange Format (LKIF) establishes an ontology and information exchange rules that support SEC regulations, forms, submissions and responses. The Financial Regulation Ontology (FRO) is an open-source ontology based on FIBO and LKIF, and consolidates regulations for banking, insurance, funds and hedge funds.

Regulatory Bodies Worldwide Are Evolving Financial Regulation Ontologies



Early adopters are utilizing the clarity and flexibility of graph modeling to create an enterprise platform for visualizing, analyzing, reporting and governing financial risk.

Neo4j customers have developed innovative solutions for addressing the data lineage and metadata management challenges of financial risk reporting.

Real-World Risk Reporting Solutions

[Neo4j customers](#) have developed innovative graph solutions for addressing the data lineage and metadata management challenges of financial risk reporting.

Online Data Distribution and Knowledge Base Platform

A leading [Global 500 financial services firm](#) needed a data distribution platform that:

- Included a knowledge base that described the lineage of datasets and attributes to online customers.
- Could accommodate new data sources, datasets, consumers and rules easily.

The knowledge base needed to be able to easily answer numerous questions such as:

- What datasets and attributes do we provide?
- How are the datasets related?
- Which consumers are using which attributes?
- How are users receiving our data?

The firm chose to store the metadata model for the knowledge base in Neo4j instead of Oracle. Neo4j's flexible schema enabled the firm to model all its data flows and rapidly answer questions about how and where its data is used.

Given the success realized with Neo4j, the firm plans on widening its coverage of datasets and offering the solution to other parts of the bank.

Real-Time Risk Assessment of Breaking News

Another leading global financial services firm is using Neo4j to make connections between news events and stock market volatility. Their system predicts potential impacts of news events on stock prices and associated portfolio exposure.

Designers of the system sought to answer questions like how a mine explosion that spikes copper prices might affect the stock price of Apple, a heavy user of the metal. The application sends real-time alerts to the phones of money and risk managers who hold Apple shares.

"These are people who don't want to be caught off-guard," says a senior vice president at the bank's innovation center. The unit contributes a few percent of the bank's revenue. But bank executives expect that revenue contribution to grow and that the new services can add value even when the bank doesn't charge for them.

"It is my hope that as we digitize the bank we will be more valuable to our customers and we will gain market share," the CEO said.

BCBS 239 and graph technology are game changers that simplify the challenges of financial risk reporting and regulation.

BCBS 239 and Graph Technology Are Game Changers

The risk-reporting mandates of BCBS 239 place new demands on data architectures at banks and financial houses worldwide. The need for fast access to real-time data lineage and financial risk information has given organizations solid justification for revisiting the old, relational reporting systems they've struggled with for years.

Forward-looking banks are using BCBS 239 as justification to proactively build federated databases that add centralized metadata control over operational data in existing silos and sources. The new risk-reporting platforms combine graph database technology and financial-industry ontologies to address crucial data lineage, latency, visualization and reporting challenges presented by BCBS 239 requirements.

As a result of their proactive efforts, innovative banks that adopt graph-based, financial risk reporting distance themselves from competition as they:

- Analyze financial risk faster and with more accuracy
- Use financial ontologies to simplify communications and speed development
- Build risk reporting applications that bring them into regulatory compliance

A Foundation for Innovative Applications

By using Neo4j to address BCBS 239 financial risk reporting requirements, you build a connected data foundation that supports a world of innovative uses of your enterprise data—including [360-degree visibility of your customers](#), [detecting and preventing fraud](#), proactively assessing credit applications, and driving what-if scenarios that improve productivity and profitability.

To learn more about how graph technology can revolutionize your financial risk reporting, or for an in-depth conversation about your specific requirements, contact Nav Mathur at nav@neo4j.com.

Neo4j is the leader in graph database technology. As the world's most widely deployed graph database, we help global brands – including [Comcast](#), [NASA](#), [UBS](#), and [Volvo Cars](#) – to reveal and predict how people, processes and systems are interrelated.

Using this relationships-first approach, applications built with Neo4j tackle connected data challenges such as [analytics and artificial intelligence](#), [fraud detection](#), [real-time recommendations](#), and [knowledge graphs](#). Find out more at neo4j.com.

Questions about Neo4j?

Contact us around the globe:
info@neo4j.com
neo4j.com/contact-us