Meeting BCBS-239 with a Risk Aggregation Platform
# Table of Contents

**Introduction** .................................................................................................................. 4
Lack of Transparency and Lax Risk Management ......................................................... 4
Reactions: Deepening and Widening Regulatory Oversight ........................................ 4
Benefits of a Strategic Approach to Compliance ......................................................... 5
Think Strategic, Act Tactical ......................................................................................... 6

**BCBS 239 – Principles for effective risk aggregation and reporting** ................. 7
A Framework for Compliance ..................................................................................... 7
Holistic Governance Framework ............................................................................... 8
Governance Competency Centre ............................................................................ 9
Governance Development Process .......................................................................... 10

**Risk Aggregation Platform Definition** .................................................................. 11
Documentation Principles ......................................................................................... 12
Metadata Driven, Active Documentation ................................................................ 12
Risk Infrastructure ..................................................................................................... 13
Data Quality ................................................................................................................ 13

**Risk Data Aggregation Capabilities** ..................................................................... 15
Aggregation Logic ........................................................................................................ 16
Aggregation Data Model ............................................................................................. 17
Virtual Aggregation ..................................................................................................... 19
Physical Data Aggregation ......................................................................................... 19
Data Persistence and Computation Infrastructure .................................................. 20

**Risk Reporting** ........................................................................................................ 20

**Conclusions** ............................................................................................................ 22

**Appendix I – Glossary of terms** ......................................................................... 23

**Appendix II – References** ....................................................................................... 23
Introduction

The Global Financial Crisis of 2007–2008, sparked off by the bursting of the sub-prime mortgage market in the US, has played a significant role in the failure of key businesses, has caused the decline in consumer wealth estimated in trillions of U.S. dollars and triggered a downturn in economic activity that has lead to the 2008–2012 global recession and contributed to the European sovereign-debt crisis.

The crisis started when the value of the US real estate market plummeted and impacted the value of securities tied to it, consequently affecting the stability of many financial institutions with exposure to a range of financial instruments linked to mortgage backed instruments.

The initial financial shocks where not quickly quelled as it became clear that the complexity of the securities vehicles and the opaque relationships between them would render the task of determining actual exposures and of unwinding positions a lengthy process.

Lack of transparency and access to timely and accurate exposure and risk information quickly led to a deepening of the crisis, with tightening of credit and liquidity affecting initially the Financial Services sector (and resulting with the collapse of many long established businesses) and quickly transmitting to the wider economy, with widespread loss of confidence causing global stock markets losses, consumer spending contraction and requiring unprecedented interventions and bail-outs from central banks, national governments and international institutions.

Lack of Transparency and Lax Risk Management

Many papers, books and reports have been produced over the years since 2007/08, documenting the general market trends and specific industry factors that contributed first to the un-restrained growth of the speculative real-estate bubble, then to its collapse and the difficulties in managing the systemic consequences. Congress and Parliamentary committees have investigated the dynamics and generally agreed that the crisis bears many parents: financial institutions, market regulators, rating agencies and governments themselves all had a hand in either directly producing, omitting to constrain or inadequately respond to the challenges.

The U.S. Senate’s Levin–Coburn Report concluded that the crisis was the result of "high risk, complex financial products; undisclosed conflicts of interest; the failure of regulators, the credit rating agencies, and the market itself to rein in the excesses of Wall Street".

The Financial Crisis Inquiry Commission concluded that the financial crisis was avoidable and was caused by "widespread failures in financial regulation and supervision," "dramatic failures of corporate governance and risk management at many systemically important financial institutions," "a combination of excessive borrowing, risky investments, and lack of transparency".

The G20 leaders declared that "weak underwriting standards, unsound risk management practices, increasingly complex and opaque financial products, and consequent excessive leverage combined to create vulnerabilities in the system", adding that "policy-makers, regulators and supervisors, in some advanced countries, did not adequately appreciate and address the risks building up in financial markets".

Reactions: Deepening and Widening Regulatory Oversight

It is beyond the scope of this paper to examine the macro-economic, structural and fiscal reforms that emerged as a reaction to the crisis, we rather want to focus on the response engendered within the global regulatory bodies and the pressure that supervisory activity has and will increasingly exercise on financial institutions.
As we’ve hinted above, some of the causes for the genesis of the crisis and certainly the difficulties experienced in the handling the stress of market bodies have been widely attributed to a lack of transparency in the composition of complex financial instruments (especially in the Over The Counter, derivatives market) and their underlying exposure to elements of risk. Risk management practices were lax and not properly accounting for concentrations of risk; furthermore, the management of Credit and Market risk was not appropriately related to the measurement and management of Liquidity Risk.

Policy makers and regulators in the US as well as international supervisory organisations have addressed these concerns through a range of new legislation, ranging from the Dodd-Frank Wall Street Act, arguably the most comprehensive market reform bill since the Great Depression, to Basel III, seeking to strengthen the banking sector through enhanced capital adequacy, stress testing and liquidity management provisions.

The Basel Committee on Banking Supervision has laid out the principles for effective risk data aggregation and risk reporting in its BCBS 239 report.

The impact of these strengthened regulatory regimes on financial institution’s businesses cannot be understressed, most often requiring a review of business models, operating capabilities and business practices.

And in this context information systems, which embody the organ supplying the data lifeblood supporting the functioning of financial firms, are being reviewed with renewed focus; the regulatory challenges to institutions are being reflected into challenges to information and data practices within the firm’s IT organisations.

In particular, BCBS 239 is very explicit in calling out the need for accurate, comprehensive, flexible data architectures to enable enhanced aggregate risk reporting.

The starting point for this request for change is the recognition that many banks’ information and data systems were inadequate to support the management of financial risk, with many banks unable to accurately and quickly aggregate risk at group level, across lines of business or legal entities.

**Benefits of a Strategic Approach to Compliance**

BCBS-239 is not the only directive that asks for radical change in the way enterprise data is managed within firms; in fact numerous regulations in many jurisdictions are directing firms towards strategic reviews, automation and holistic governance of data flows; examples include Dodd-Frank, EMIR, SEC CAT, to name but a few. The common denominator of all these regulations is that supervisors want better visibility of data and information as they move within the organisation and between firms.

Adopting an enterprise-wide and strategic approach to data management would enable banks to implement compliance to multiple regulatory regimes much more flexibly; a well architectued data governance framework, supported by appropriate tools and processes will enable firms to achieve compliance with new requirements through incremental deliveries built on a consolidated “compliance plant”, rather than have to spin off an entire programme with associated software, infrastructure, systems development and delivery lifecycle.

In effect, we would argue that “compliance” would be realised as a side benefit of implementing a holistic data management and governance framework. The cost benefits of such an approach are quite simple to establish and a number of industry studies by research firms such as Gartner, Forrester and AMR Research can be helpful to highlight the benefits; a Gartner study suggests a 30% of integration projects savings can be achieved through the establishment of an Integration Competency Centre. A detailed discussion of savings and benefits that can be derived from the establishment of a factory approach to integration can be found in John Schmidt’s book on Lean Integration.

McKinsey estimates banks could save between 5-10% on of some of their big operational cost categories through “better data management and reduced number of reconciliation.”
Another benefit firms stand to achieve is increased business agility: visibility and transparency of data would lead to better decision-making, both at strategic and at business level.

But arguably the most valuable upshots yet will be derived from the analytical exploitation of integrated data delivered by the compliance-driven programmes; specifically considering the Risk domain, accurate, complete and timely aggregate views of risk will enable firms to optimise their portfolio strategies and better allocate their economic capital. In the same report (1), McKinsey point out that “one bank, for example, was able to eliminate its $7 billion Financial Services Authority–mandated liquidity buffer, which carried an estimated 4 per cent opportunity cost”.

**Think Strategic, Act Tactical**

For all these reasons, we believe that firms should think strategically about how they manage data. But our longstanding involvement with firms in some of the most far-reaching, transformational information management programmes teaches us that the best advice we could give is for firms to think about tactical deliveries. It may seem at first a contradiction in terms, but through this paper we will show that it is possible to reconcile the two: a strong Compliance (and more generally, Information Management) framework, supported by an appropriately architected Platform, can deliver against the strategic goals through bite-size, incremental deliveries, each contributing incremental business value.
BCBS 239 – Principles for effective risk aggregation and reporting

Following on from previous guidance, the Basel Committee has clarified a set of principles, which systemically important banks should implement with respect to the management of risk information.

Adoption of these principles will ensure that risk aggregation and risk reporting capabilities are enhanced to support improved risk management and decision making within firms, with better reporting and strategic decision making; better visibility of risk across legal entities and business units will contribute to reduced probability and severity of losses, as the speed at which decisions can be made will be improved.

The principles cover four related topics:

- Overarching governance and infrastructure
- Risk data aggregation capabilities
- Risk reporting practices
- Supervisory review, tools and cooperation

The risk aggregation capabilities and reporting practices have been considered separately, with sound governance and infrastructure ensuring that information flow from one to the other. As some of the principles may give rise to conflicts (e.g. the need for accuracy and frequency), it will be responsibility of the firms in collaboration with supervisors to ensure the appropriate compliance to the principles is implemented.

A Framework for Compliance

In line with the typical governance approach adopted by European and International supervisory organisations, the Basel Committee lays out the requirements for the management of Risk Aggregation and Reporting processes through principles-based guidance. The principles lay out the requirements for the establishment of a governance framework and describe some of the capabilities that Banks will need to implement to support high quality risk management processes, supporting both supervisory compliance and internal strategic planning and management reporting needs.

The principles are categorised as in the following table:

<table>
<thead>
<tr>
<th>Overarching Governance and Infrastructure</th>
<th>Risk Data Aggregation Capabilities</th>
<th>Risk Reporting Practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Governance</td>
<td>- Accuracy and Integrity</td>
<td>- Accuracy</td>
</tr>
<tr>
<td></td>
<td>- Completeness</td>
<td>- Completeness</td>
</tr>
<tr>
<td></td>
<td>- Timeliness</td>
<td>- Timeliness</td>
</tr>
<tr>
<td></td>
<td>- Adaptability</td>
<td>- Adaptability</td>
</tr>
<tr>
<td></td>
<td>- Data Architecture and IT Infrastructure</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Accuracy</td>
<td>- Clarity and Usefulness</td>
</tr>
<tr>
<td></td>
<td>- Comprehensiveness</td>
<td>- Frequency</td>
</tr>
<tr>
<td></td>
<td>- Clarity and Usefulness</td>
<td>- Distribution</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Supervisory review, tools and cooperation</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Review</td>
</tr>
<tr>
<td>- Remedial action and supervisory measures</td>
</tr>
<tr>
<td>- Home/Host collaboration</td>
</tr>
</tbody>
</table>

Table 1 - Risk Aggregation and Reporting Principles
BCBS 239 makes it clear that all the principles should be met simultaneously, even when the requirements of one may be in contrast with another; it does indeed recognise that trade-offs may become necessary in exceptional circumstances, for example when responding to urgent/ad-hoc request for information, but requires that such trade-offs have no material impact on the decision making process, that appropriate policies and processes are in place and that banks are able to fully explain them.

The principles are explicit in calling out the need for an overarching and comprehensive framework for compliance. They also spell out the characteristics and capabilities that should be implemented by the framework and that will enable banks to fulfil the requirements for risk aggregation and reporting, both in normal times, but also during times of stress or crisis.

It is of crucial importance that banks consider the completeness and consistency of the framework adopted to implement compliance to the Principles; given the inter-dependencies between Principles, the variety of internal and external actors involved, the constant evolution of the institutions and market change we would advise banks to reject the temptation to create an evolutionary framework based on current policies and systems and rather exploit the opportunity to adopt a state of the art Risk Aggregation Platform, built to embed the Principles at the core of its Architecture.

Adoption of a Platform approach will ensure that the compliance framework meets the supervisor’s requirements not only at point in time, but will deliver a set of capabilities that will be applied flexibly and consistently as the business architecture evolves.

**Holistic Governance Framework**

BCBS 239 Principles should be recognised by firms not only as a legal requirement to meet supervisory compliance, but as an opportunity to embed best practice processes, tools and capabilities to manage their most valuable enterprise asset: data.

The principles ask for the establishment of holistic governance of risk data and most firms have already implemented some of the elements supporting a compliance framework, for example through the establishment of competency centres or centre of excellence in some of the key disciplines required to implement the Governance framework: typical competency centres have been created to focus on Integration (Data or Application), Business Intelligence or indeed Data Quality.

Most firms have also selected tools to support these competency centres and enterprise standards have been set and in some cases adhered to.

An effective Governance framework rests on the ability to define valid processes, supported by adequate technologies that can be applied consistently across the bank; the appropriate Target Operating Model needs to be defined and suitable management structures put in place to support it.

What has been missing so far has been the Board level mandate to define a compliance-enabling Operating Model, which will unify processes, define enterprise-wide technology standard and establish oversight and measurement of implementation best practices.

BCBS 239 and a number of national and supra national regulatory initiatives accompanying it are pushing banks to establish new executive roles, often with Board level responsibility, to manage Data within this new Regulatory regime: Chief Data Officers and Chief Data Architects are now working closely
with their colleagues at the head of Governance, Risk and Compliance (GRC), Enterprise Architecture and Information Technology to establish and embed the processes, tools and skills required by the governance regime.

**Governance Competency Centre**

A useful construct for the implementation of a strong and adaptable Governance framework can be borrowed and built on from the IT practice of establishing an Integration Competency Centre (ICC) to improve the quality and reduce the cost of running multiple data integration projects.

The basic ideas behind the ICC practice are simple; it is the embodiment of IT management best practices to deliver shared services. The same concepts can be applied to the establishment of a Governance Competency Centre, especially in consideration of the similarities of processes, tools and deliverables that are the subject matter under its responsibility. The goals of the Governance Competency Centre should be:

- Lead and support data and information Governance (data, system and process) projects with the cooperation/coordination of subject matter experts
- Promote Governance as a formal discipline. For example, Governance will include data stewardship, modelling, aggregation, documentation and reporting, disciplines.
- Develop staff specialists in Governance processes and operations and leverage their expertise across multiple compliance initiatives.
- Manage data aggregation and reporting pilots and projects across the organization
- Optimize investments (skills, tools) across multiple projects and leverage economies of scale.

A Governance Competency Centre will enable the bank to:

- Optimise scarce resources by combining Regulatory, Risk, Trading and Technology skills, resources, and processes into one group
- Reduce regulatory compliance delivery times and development and maintenance costs through effectiveness and efficiency
- Decrease duplication of compliance related effort across multiple supervisory regimes through creation and maintenance of comprehensive and integrated view of data and process assets
- Improve ROI through creation and reuse of enterprise assets like business glossary, data models and maps, application interfaces, and well documented business rules
- Build on past successes instead of reinventing the wheel with each project
- Lower total technology cost of ownership by leveraging technology investments across multiple compliance projects.
In order to meet the requirements expressed by the Principle, but also to ensure that the Governance Framework implemented by the bank is flexible enough to respond to additional requirements and supervisory challenges, the Governance Process established should favour Agile process patterns; specifically, it should support iterative phases, allowing the outputs from one phase to be continuously fed as input into the other phases.

As mentioned earlier, it is of crucial importance that the overall Governance Framework contributes value to the organisation through phased deliverables. The iterative approach of the Development Process will enable such deliverables; in order to ensure the capabilities implemented at each phase deliver incremental value and align to overall objectives, a framework should be established to assess, measure and monitor progress through the phases of the cycle and over multiple iterations.

**Figure 2 - Development Process**

- **Documentation** - At the heart of the Governance Process should therefore lie a Documentation phase, which will support, drive and enable all the other activities and phases of the process. Documentation will include the measurement of activities, providing visibility in attainment of Compliance and wider Business Value generation.

- **Discovery** – Activities in this phase are concerned with the assimilation of data assets into the Governance framework: systems and data items definitions are imported and appropriate Documentation is created to describe their topology, semantics and taxonomies. Typically, a Business Glossary of terms is established, documenting both industry and Bank terms of reference. A complete list of all data assets is created and published.

- **Analysis** – in this phase the initial list of data assets is augmented with in-depth documentation of data types and values. These are resulting from inspection of actual data, through techniques of data sampling and deep-profiling; any discrepancies between expected (e.g. from pre-existing data models, database schemas, etc.) and actual values are flagged and highlighted for consideration, as appropriate measures will need to be put in place to integrate and reconcile discrepancies. Data elements are fully documented, including structure (e.g. data elements, their types and data domains), function (keys, measures, etc.) and quality details, including accuracy, integrity and completeness.
• **Model** – Data definitions, taxonomies and profiles will be enriched with relationship definitions, establishing links between entities, establishing entity owners, relating data to their consumers; the documentation redacted at this point of the process should include data lineage reports, linking source to consumer and each data entity to their data profile. A complete model (either fully integrated, or federated set) of Bank’s risk is the result of this phase.

• **Rules Definitions** – Once a common taxonomy of terms have been defined, data entities have been documented and relationships between them established, data owners will define the rules that govern the aggregation of data to compose complete views of risk. These rules definitions, as authored by the business owners, will provide a functioning prototype for the implementation of the aggregation rules into a set of mapping logic or templates, which will be transformed into operational integration by Risk IT. It should be noted that the handshake between business owners (front office trading, risk management, controlling functions) and Risk IT staff should be facilitated by the utilisation of common language and definitions (business glossary and taxonomies) and shared tools to prototype and define data aggregation logic, e.g. high level business rules editing tools, data flow diagrams, etc.

• **Data Aggregation Flows Implementation** – the specifications for the aggregation of risk data, as defined by the business users and data owners, will be used by Risk IT to implement production-ready data aggregation flows. A high level of abstraction between rule definition and physical execution should be sought in order to maintain the ownership of the rules definition as close as possible to the business owners, while allowing Risk IT to optimise the flow of data to meet timeliness and frequency (throughput and latency) requirements.

### Risk Aggregation Platform Definition

The primary characteristic and cornerstone of the platform’s architecture should be its metadata foundation, which will underlie all the capabilities provided and support the functional components implementing the capabilities required to meet the Principles.

Metadata will also abstract the functional capabilities of the Platform from the underlying Infrastructure, ensuring that the functional components of the platform are underpinned and enabled by a scalable and highly available infrastructure foundation.

Principle 2 of BCBS 239 very explicitly demands that the risk data aggregation and reporting practices should be supported in normal times as well as in periods of stress or crisis.

This principle specifically aims to address the difficulties experienced by SIFIs during the summers of 2008 and 2011, when the volume of trades at the height of the mortgage back credit and sovereign debt crisis more than doubled in volume.
Documentation Principles

Documentation should be created and maintained to represent the state of the overall Architecture, and of the processes that support the firm’s Risk Aggregation Capabilities and Reporting practices.

- **Independent Validation** - Principle 1 of BCBS-239 requires that the capabilities be independently validated and reviewed against the principles. It is therefore of primary importance that documentation provides an accurate reflection of the capabilities and practices; it should also be all encompassing, providing visibility into all the components of the architecture and all the processes that is support.

- **Transparency** - Documentation should be easily inspected by both internal reviewers and external supervisory authority staff; this will enhance transparency and reduce the costs associated with review and audit processes. Transparent documentation will also allow the identification of gaps and opportunities for improvement of the capabilities.

- **Comprehensive documentation** – a Risk Aggregation Platform where all components share a common Metadata foundation will generate documentation automatically, linking the definitions of terms, data items, the data profiles, the relationships between data and information objects, the data transformation logic (validation and integration) and the reporting structures. The same Metadata foundation should also be extensible to allow the incorporation of additional documentation, including specifications, data models, process definitions

- **Accuracy of documentation** – The platform will facilitate the production of documentation, as this is simply a representation of the metadata assets that define the data aggregation and reporting processes. Every item of data and function applying to them are defined through metadata objects; as such, documentation is always “active”, meaning that it reflects the present characteristics of data, information and processes. Automatically generated metadata can be navigated by users via appropriate interfaces; these enable users to search, catalogue and extend the documentation; metadata tags, definitions and data quality profiles can also be incorporated in reports and dashboards for management information distribution.

Metadata Driven, Active Documentation

Metadata should be automatically created throughout the data aggregation and reporting process and should document and transparently expose the risk aggregation capabilities and reporting practices through appropriate and tailored mechanisms. Tools should be provided to flexibly extend, integrate, steward and curate automatically generated metadata, which should include the following classes of information:

- **Data Ontologies** – establishing common taxonomies and glossaries of terms across the bank, defining the identifiers, counterparties, legal entities, customers and accounts.

- **Data Quality** – metrics capturing the quality of data against a comprehensive set of dimensions should be automatically captured; reconciliation of data across validation points should also be captured as required.

- **Data Semantics** – extending the data classifications with relationships between entities and classes of information. Visibility of relationships should be transparent to Business and IT resources; the depth of the Semantic layer should extend to relate data with their Ontologies as well as their Quality measurements.

- **Aggregation Logic** – rules for the aggregation of data across lines of business, geographies, legal entities etc. should be defined in collaboration between Business data owners and IT; of course their implementation should directly reflect their definitions, but it should also be exposed transparently, so that inspection is facilitated. Processing logic defined through metadata should also enable the profiling of the rules, in order to drive continuous improvement in data aggregation capabilities.
• Security and Access – information should be made available to users according to their roles and responsibilities and access to information should be audited.

• Service Levels and Operations - information about the operational availability of the systems and processes should be recorded and monitored and appropriate alerting and suitable processes executed to maintain or restore service levels.

• Compliance and Business Value – metrics measuring the progress towards compliance goals and delivery of wider business value to the organisation should be provided as part of the documentation body.

It should be noted that documentation should be provided by the platform as a set of “Active” artefacts, reflecting a view of content stored in the shared metadata repository, rather than be implemented as a post-facto exercise resulting a point-in-time snapshot of the architecture.

The active nature of metadata will facilitate the sharing of documentation between all the stakeholders involved in the governance framework, both internally to the bank (GRC and line of business functions) and external regulatory bodies.

Risk Infrastructure

Risk Aggregation and Reporting Infrastructure should benefit from the developments in technical architectures experienced over the last few years, for example, the improvement in data processing capacity provided by grid and scale-out architectures in both Relational and NoSQL forms.

• Scalability - A platform supported by grid computing infrastructure demonstrating linear scalability across compute nodes provides the mechanism to predictably manage both expected data volume growth and respond to increases in transaction rates in period of stress.

• Availability - Grid computing also provides the means to implement high availability, by distributing workload across a set of resources with spare capacity and fail-over capabilities it ensures that processing can be performed even in the event of failure of some of the components.

• Optimisation – specific functional tasks, for example the computation of aggregate risk measures at portfolio level, correlation of risk factors, etc. may require the adoption of specific optimisation techniques, including in-memory data processing, or the statistical scoring of records; the platform should ensure that the Process is taken to the data (push-down), minimising the movement of large volumes of data across the components of the Platform.

• Security – Data should be subjected to stringent access and authorisation controls as governed by the access policies, throughout their lifecycle and irrespective of the storage or persistence medium on which they reside. Data should be encrypted when in transit as well as when at rest and made available exclusively to the authorised users.

Data Quality Risks

The platform should provide holistic capabilities to manage the quality of data supporting both in house and outsourced risk management processes and should enable the bank and the supervisor to apply suitable governance oversight over the data and reporting lifecycle (see §27 of the Principles paper).

These capabilities apply to the extant data and are delivered through a unified framework that facilitates for the on-boarding and management of new data created or integrated as a response to new business development initiatives.

• Identification of data, this should include the automated discovery of relevant attributes and relationships between data entities and most critically ownership and responsibility for the stewardship of data is established. Identification of gaps in data quality should be used to define the processes for the validation and correction of said data; relationships identified should
be utilised to enrich the data model and provide a more comprehensive view of risk; where the relationships between entities require complex mapping to resolve structural inconsistencies (e.g. between lines of business) the framework will provide the capability to easily establish a process for the appropriate mastering of cross reference data.

- **Assessment** of quality metrics should be performed against a range of dimensions, and BCBS calls out specifically some of them: accuracy, timeliness, completeness, etc. The framework should implement automated measurement of all relevant data against these metrics and provide data quality scoring, which should accompany the data. Data quality metrics should therefore be available for inspection at any level of aggregation (at individual trade level, legal entity, business unit or group) and retained over time to allow for monitoring of trends and to drive continuous improvement.

- **Monitoring** of data quality risks should enable “management by exception” processes to be implemented, ensuring that activity is targeted and focused on automation improvement and process optimisation; to this aim measurements against the metrics defined during the assessment phase need to be captured at regular intervals and exception alerts executed as required.

- **Management** of data quality risk requires the establishment of appropriate policies and processes, which will govern the resolution of gaps identified; these may involve the adjustment of data or source systems under appropriate change management processes or the implementation of automated and audited reference data mapping through the aggregation process.
Risk Data Aggregation Capabilities

The principles related to risk data aggregation call for the processes applied to be largely automated as to minimise the probability of errors, assuring the accuracy and integrity of data. Manual processes should be mitigated by appropriate policies and procedures.

At the same time, BCBS requires that the risk data should be captured and aggregated across the entire banking group, meeting a completeness principle.

Recognising that banking business is in continuous evolution, BCBS asks that the data aggregation architecture is flexible and adaptable to meet changing internal needs and supervisory requests.

The fourth principle, that of timeliness, requires that risk information be available regularly, frequently and in a timely manner.

Clearly BCBS recognises, perhaps in consideration of the typically siloed nature of risk systems in most SIFIs, that these principles are somehow at odds with one another: preparation of data should be highly automated, industrialised and standardised, yet flexible enough to cope not only with planned business evolution but also with responses to ad-hoc regulatory requests. Data should be accurate and surrounded by controls on a par to accounting data, and yet up-to-date at all aggregation levels.

The Risk Aggregation Platform will meet the requirements outlined through these principles by unifying what were until recently very distinct data integration management methodologies: ETL processes will persist data where appropriate and the platform’s Virtualisation capabilities will extend and complement persisted data to meet the agility or timeliness requirements as required.

It is fundamental that the both integration paradigms share a common definition of the bank’s data assets (taxonomies) and implement the same model through consistent validation and aggregation rules.

The choice between physical and virtual realisation of data should not be dictated a-priori by tool philosophy or choice, but rather as a function of the timeliness and agility requirements.
Aggregation Logic

As discussed above (Documentation Principles)

Documentation should be created and maintained to represent the state of the overall Architecture, and of the processes that support the firm’s Risk Aggregation Capabilities and Reporting practices.

- **Independent Validation** - Principle 1 of BCBS-239 requires that the capabilities be independently validated and reviewed against the principles. It is therefore of primary importance that documentation provides an accurate reflection of the capabilities and practices; it should also be all encompassing, providing visibility into all the components of the architecture and all the processes that is support.

- **Transparency** - Documentation should be easily inspected by both internal reviewers and external supervisory authority staff; this will enhance transparency and reduce the costs associated with review and audit processes. Transparent documentation will also allow the identification of gaps and opportunities for improvement of the capabilities.

- **Comprehensive documentation** – a Risk Aggregation Platform where all components share a common Metadata foundation will generate documentation automatically, linking the definitions of terms, data items, the data profiles, the relationships between data and information objects, the data transformation logic (validation and integration) and the reporting structures. The same Metadata foundation should also be extensible to allow the incorporation of additional documentation, including specifications, data models, process definitions

- **Accuracy of documentation** – The platform will facilitate the production of documentation, as this is simply a representation of the metadata assets that define the data aggregation and reporting processes. Every item of data and function applying to them are defined through metadata objects; as such, documentation is always “active”, meaning that it reflects the present characteristics of data, information and processes. Automatically generated metadata can be navigated by users via appropriate interfaces; these enable users to search, catalogue and extend the documentation; metadata tags, definitions and data quality profiles can also be incorporated in reports and dashboards for management information distribution.

section), aggregation logic should be defined in a metadata repository that will be shared across the functional components of the platform.

Aggregation rules will define the data validation, merging and transformation logic that enable the bank to consolidate and integrate risk data originated from the enterprise silos. Definition of the aggregation rules will be facilitated by the metadata captured through the process of profiling the data and will be stored as shared metadata assets, enabling standardisation and re-use to improve quality and reduce implementation times.

It has to be recognised that most banks have distributed risk calculation rules across a plethora of legacy systems and platforms, spanning organisation boundaries (business and IT), architecture domains (DWs, Analytics, dedicated computation engines) and programming languages (everything from Java to Excel calculations and macros, with anything in between, including database languages, BI and Reporting tools, etc.).

One of the primary objective of the Risk Aggregation Platform should be to provide a governance framework to facilitate the:

- Cataloguing and documenting of this multitude of rules: this should be conducted through an automated discovery exercise, capturing as much as possible of the rules logic into the shared Metadata Repository. When automated harvesting of rules should not be possible, data interfaces should be full documented and rule code logic captured and integrated in documentary (source code) format within the metadata repository.

- Consolidation of the rules to a consistent set of validated and certified definitions of such rules
• Establishing the most appropriate tiers of the architecture where Risk calculations should be performed, and ensuring that appropriate mechanisms are in place to transfer rules from business-user managed tools (e.g. visualisation, business intelligence, spread-sheets) into fully managed environments as appropriate.

Data profiling will provide inferred data type definitions, data domain membership and establish relations between data entities, which will enable the definition of specifications or templates for the implementation of aggregation rules.

The Aggregation Platform provides mechanisms for the consumption and processing of data in the appropriate unit of work: exposures and metrics that can be made available in real-time will be processed and made available for reporting as soon as available; measures requiring set based processing (e.g. due to correlations) will be handled as appropriate and data timeliness indicators will be available to the consumers as additional data quality metrics.

The platform will therefore support a range of data exchange mechanisms, including publish/subscribe methods, event driven, on-demand and scheduled data transfer and processing; the specific characteristics of the Risk data provider and requirements of the consumer will determine which mechanism shall be utilised for exchanges. The platform will implement the appropriate mechanisms for the timely supply of data aggregated at the required levels of granularity; these may include a range of optimisation techniques and appropriate infrastructure configurations, for example:

• Event-driven computation of risk – utilising in-memory or event processing paradigms to pre-compute risk measures aggregate at specific levels of granularity (e.g. business, geography, counter-party legal entity, markets, etc.)
• Summary / aggregate database views – periodically refreshed to provide suitable response time to ad-hoc user access or dashboard updates
• OLAP technologies, allowing periodic / on demand recalculation of multi-dimensional storage models

See the “Data Persistence and Computation Infrastructure” section for more information on some the technologies available and relevant techniques applicable to the domain.

Aggregation Data Model

One of the biggest challenges in delivering group-wide representation of Risk data has historically been the difficulty in integrating and unifying a data model to represent all types of risk data across business units, legal entities and geographies.

In fact, BCBS-239 does not mandate the creation of a single data model, but asks that “robust automated reconciliation procedure” be implemented when multiple models are in use.

The starting point for the construction of a comprehensive “aggregate” view of risk, as we’ve pointed out, should start with discovery of existing assets; these include extant reporting applications, risk calculation systems and reference data providers.

Special consideration should be given to reference and master data, as these will invariably provide the “keys” allowing the merging and aggregation of disparate transactional level data; it is often the lack of strategy or implementation resolve of Master Data Management capabilities which hampers strategic data

Figure 5 – Logical Data Model
management programmes, including Single View of X (Customer, Product, etc.) or Enterprise Data Warehouses.

Failure to establish “master records” for key entities such as Counterparty, Instrument, Book, etc. prevents the resolution of silo data sets into data super-sets, requiring extensive reconciliation processes with manual intervention for record allocation.

In order to enable the creation of an aggregate view of Risk, it is fundamental that the Risk Aggregation Platform should provide tools to manage master and reference data and specifically equips the Risk Aggregation program with capabilities to:

- Define master record attributes for both simple and complex entities (e.g. Party, Instrument): banks should look for vendor or market accelerators in the form of reference data models for such entities if possible, but we would strongly encourage banks to favour model flexibility over off-the-shelf model completeness.
- Define and maintain multiple and evolving hierarchies of these entities; the reference and master data model need to reflect changes in Business hierarchies, Customer Legal Entities, and these changes need to be visible and reflected both in any reporting and analytical applications and operational systems.
- Match and resolve records into the master entities; this will require extensive and flexible matching rules to be defined and maintained; business users should be provided with easy to use interfaces allowing them to define the rules that determine record or attribute survivorship; sophisticated fuzzy logic matching should be available to match textual and complex set of attributes.
- Facilitate the joint ownership of the solution between Business and IT; the model and rules are owned by Business and the tools need to seamlessly extend to allow robust integration of reference data into the Risk Aggregation Platform; in fact, the platform should implement a consistent model-driven paradigm to both master data management and data integration.

Best practice indicates that adoption of an agile, iterative process for the delivery of a comprehensive model is to be favoured over a sequential, waterfall-style implementation, especially in consideration of the capabilities of the Platform to support close integration of Risk users and Risk IT working groups through shared documentation and active review of data at any stage of the process of integrating the model.

The Data Aggregation Platform will facilitate the creation of a unified model and the implementation of appropriate reconciliation processes should ahead of full unification: the shared metadata capabilities of the platform will allow the importation of existing models (from enterprise modelling tools, reverse engineering existing database or spreadsheet) and their integration and evolution through rapid prototyping techniques.

Rapid prototyping should be used as a means of validating design in view of creating a conform model and it is significantly empowered by the ability to surface model objects through the Data Virtualisation capabilities of the platform.

Rapid Prototyping is also a powerful technique that enables the bank to quickly respond to business initiatives such as acquisitions, divestitures or product introduction and should be used extensively to validate the evolving model definition internally with front office staff, risk management and control functions and externally with the Supervisory authorities.

Figure 6 Reconciliation Reports
Where multiple definitions of entities exist, the Platform will provide the facilities to easily create reconciliation reports; robust reconciliation processes will ensure manual as well as automated supervision of data aggregation.

**Virtual Aggregation**

Data objects represented within the Metadata Model can be implemented as Virtual representations of the data; this technique, sometimes also referred to as “Data Federation” doesn’t require the persistence of the data into a physical storage medium (such as a database) rather resolves the data entity whenever the object is requested from a Report or queried by a User.

Virtual Aggregation offers the bank the ability to dramatically reduce the typical software development lifecycle (SDLC) time-span required to integrate new data into the unified Aggregate Risk Model, therefore meeting the requirements as expressed in Principle 6 of BCBS 239.

Virtualisation also provides another capability worth of notice: it enables the extension of the model to incorporate highly refined and certifiably accurate data (i.e. data that has been through the validation process) with more up-to-date, yet to be validated data. Again, up-to-date, but possibly less accurate, data will be tagged with appropriate data quality (accuracy or integrity) tags, ensuring that both Principle 3 and 5 are met simultaneously.

Clearly, virtualised aggregation provides agility and extensibility benefits, but these are often counter-balanced by the potential degradation of performance associated with applying complex aggregation logic requiring movement of large volumes of data at query time. The Virtual Aggregation capabilities of the platform should therefore be supported by optimisation capabilities, such as caching, that would alleviate these effects.

**Physical Data Aggregation**

Physical data aggregation has historically been implemented through a set of techniques and tools often referred to as Extract, Transform and Load (ETL).

The Aggregation Platform will implement the widest range of ETL capabilities, allowing the collection of data in the most appropriate manner depending on the provider of data and required processing mode; for example, the Platform will leverage high throughput, parallel database unload techniques to extract large sets of data or will consume events from messaging channels. Changed Data Capture techniques will be provided to “event enable” sources of data that don’t natively support event feeds (e.g. databases).

The Physical Data Aggregation component of the Platform will share the same model and aggregation rules with the Virtual Aggregation component, therefore ensuring consistency of definitions and processing.
Data Persistence and Computation Infrastructure

Aggregate risk data will be persisted in a high performance data management platform, supporting high throughput concurrent load and query workloads. In order to accommodate the requirement for event based loading of the risk data, the Persistence store will support sophisticated multi-version consistency locking mechanisms. It will also support the requirements for availability and scalability through suitable mechanisms for data distribution across storage nodes and failover to grid computing resources.

A number of technology innovations have become available over the last decade which are providing a number of highly desirable characteristics, such as:

- Distributed data storage, notably HDFS (Hadoop File System) at the NoSQL end of the spectrum, and corresponding MPP (Massively Parallel Processing) or “shared nothing” DBMSs (Database Management Systems); these systems providing scale-out architectures typically implement High Availability through fault tolerance and graceful failover and deliver linear and predictable scalability as nodes are added to the network. Commodity building blocks (typically X86 architecture based) provide a cost efficient model for delivery of high performance, reliable infrastructure resources.

- In-memory processing, where most of the typical data management operations carried out by a DBMS have been handled entirely in-memory has been enabled by the dramatic increase in memory capacity and its corresponding decrease in cost; this has been coupled with innovation by leading vendors, including SAP (HANA), Oracle, IBM and Microsoft, each with product options either generally available on the market or announced for imminent release.

- Event driven computation (Complex Event Processing) technologies have become available across distributed architectures, again supporting scalability and fault tolerance; these technologies have been utilised to aggregate risk exposure and sensitivities in response to trade events and coupled with persistence stores to provide consistent views of risk from Front to Back office.

Risk Reporting

In traditional implementations of information management systems business users interaction with the information management platform is reserved to their consumption of Business Intelligence and Reporting toolkits. What we advocate, as discussed throughout this paper, is that business users ranging from Risk Management, Risk Controllers, Heads of Trading, Senior Management and representatives of Supervisory agencies should have access to a number of tools that allow them to collaborate throughout the data gathering to exploitation lifecycle: common definitions of terms, standardised interpretation rules, relationships between entities and systems should be the result of the cooperation of business and IT.

The shared ownership and governance of enterprise Risk data assets should of course extend to reporting and analytical exploitation phases of the lifecycle.

Reporting, visualisation, analytical exploration tools provide great capabilities allowing users to play with and “model” data, adding context and representational characteristics appropriate to their intended consumers. This level of flexibility is highly desirable, but has historically come at a cost: often it has resulted in fragmentation of data and data interpretation rules, resulting in inconsistent information being presented across a range of documents.

The Risk Aggregation Platform and the Governance Framework surrounding it should make every effort to minimise this down-side while preserving flexibility and user freedom to interpret data; these efforts can be facilitated by the availability of shared, clear, accessible and up-to-date documentation about the available Risk data.

Maintaining control of definitions (of entities and rules) will require specific consideration to be given to the following aspects:
• Ownership: correct allocation of responsibilities and ownership of rules and data is of crucial importance; clarity over their allocation will help define appropriate processes for their management.

• Agile development: typically definitions will be gathered during Analysis phases of the development lifecycle, but it is also possible that changes to entities and rules will result from business user access for reporting and analytical exploitation of information.

• Tool flexibility: as discussed throughout this paper, we advocate a “documentation-centric” approach to the development lifecycle, and propose the adoption of a Platform resting on a shared, active metadata layer. It should be possible, and indeed required, for business users to own the data and the rules, and for them to discover and define them in the set of tools and user interfaces most appropriate for their tasks. It should be possible for business rules, for example the creation of compound measures of risk, to be created within Business Intelligence and Reporting tools. Shared metadata should be used to ensure that the creation of such rules is documented and highlighted, such that appropriate governance over them can be applied.
Conclusions

Banks have often approached Regulatory Compliance as a burden to be met with ad-hoc responses and through very specific and cost efficient measures. It is clear that the Financial Crisis witnessed over the last decade and the Supervisory agencies response to it has dramatically changed the game: banks are asked to respond by setting up Board level functions to oversee and Govern the creation, management and analysis of Risk data.

Banks should take this opportunity to invest strategically in comprehensive Governance frameworks that will establish Best Practice principles, processes and technologies to meet present and future challenges systematically.

This paper has provided the reader with insight into these Best Practices and some of the leading capabilities that should be demanded from an investment in technology to support the Governance framework. Specifically we have articulated the need to approach technology choices from the perspective of building an integrated, self-documenting Risk Aggregation Platform that will enable agile response to evolving Compliance requirements.

To find out more about Informatica and how its products and services can help organisations implement Governance framework supported by a Risk Aggregation Platform please visit www.informatica.com
Appendix I – Glossary of terms

**BCBS 239** – Basel Committee on Banking Supervision – Principles for effective risk data aggregation and risk reporting. The paper can be found on BIS’s website.

**GRC** – Governance, Risk and Compliance – functions within the Institution tasked with ensuring compliance with relevant laws, regulatory regimes and establishment and adherence to internal Risk Management practices.

**gSIB(s)** – global Systemically Important Bank(s) – those banks whose failure is deemed to represent a risk to the global financial services market. The list of gSIBs can be found on the Financial Stability Board website. As of November 2013, the list comprises 29 banks and will be revised again in November 2014.

**gSIFI(s)** – global Systemically Important Financial Institutions – those institutions with a global presence and whose failure may trigger a financial crisis. A list of gSIFIs can be found on the Financial Stability Board website.

**Lean Integration** – a management style that emphasises continuous improvement in the end-to-end data integration and process integration activities. More can be learned about Lean Integration in John Schmidt and David Lyle’s book

**Integration Competency Centre (ICC)** – a shared function within an enterprise organisation tasked with performing methodical data, system or application integration programmes. More information on ICCs can be found on Wikipedia

Appendix II – References

1. McKinsey & Company, 2013, It’s not just about getting your ratios right
2. Gartner Group, 2008, “Cost cutting through the Use of an Integration Competency Center or SOA Centre of Excellence”