

Implementing Data Quality in BCBS 239 Requirements: The Case of G-SIBs



Received: 18/12/2024; Accepted: 23/03/2026

Zineb REDOUANE ALI^{1*}, Mohammed DEHANE²

¹ Abdelhamid Mehri University, Constantine 2, Algeria, email: redouaneali.zineb@univ-constantine2.dz

² Abdelhamid Mehri University, Constantine 2, Algeria, email: mohammed.dehane@univ-constantine2.dz

Abstract

In light of modern developments, business complexity, the banking sector's success increasingly relies on effective data collection and analysis. High-quality data is crucial for risk assessment, customer profiling, operational efficiency, regulatory compliance, and maintaining a competitive advantage. The study relies on reports from the Bank for International Settlements (BIS) to evaluate compliance with the BCBS239 principles, with a focus on the strengths and weaknesses of (G-SIBs). It emphasizes the critical role of data quality in BCBS239 adherence, focusing on dimensions such as accuracy, integrity, adaptability, timeliness, and completeness. Governance and data infrastructure are identified as key determinants of data quality, significantly affecting risk management and regulatory reporting. The study offers actionable recommendations to improve data quality, underscoring its importance in regulatory compliance and strategic decision-making, particularly during crises. It also offers a comprehensive analysis of the role of data in enhancing resilience and informed responses in risk scenarios.

Keywords

Data quality;
BCBS 239;
Risk data aggregation;
BIS reports;
G-SIB.

الكلمات المفتاحية

جودة البيانات؛
BCBS 239
تجميع بيانات المخاطر؛
تقارير BIS؛
G-SIB.

تنفيذ جودة البيانات في متطلبات BCBS 239: دراسة حالة G-SIB

ملخص

في ظل التطورات الحديثة وتعقيد الأعمال، يعتمد نجاح القطاع المصرفي بشكل متزايد على جمع البيانات وتحليلها بفعالية. تُعد البيانات عالية الجودة ضرورية لتقييم المخاطر، وتحليل العملاء، وتحسين الكفاءة التشغيلية، والامتثال التنظيمي، والحفاظ على ميزة تنافسية. تعتمد الدراسة على تقارير صادرة عن بنك التسويات الدولية (BIS) لتقييم الامتثال لمبادئ BCBS239، مع تسليط الضوء على نقاط القوة والضعف في البنوك ذات الأهمية النظامية العالمية (G-SIBs). تُبرز الدراسة الدور الحيوي لجودة البيانات في الامتثال لمتطلبات BCBS239، مع التركيز على أبعادها الرئيسية مثل الدقة، والنزاهة، والقابلية للتكيف، والتوقيت، والشمولية. كما تُحدد الحوكمة والبنية التحتية للبيانات كمعوامل رئيسية تؤثر على جودة البيانات، مما يؤثر بشكل كبير على إدارة المخاطر والتقارير التنظيمية. تُقدم الدراسة توصيات عملية لتحسين جودة البيانات، مُشيرة إلى أهميتها في الامتثال التنظيمي واتخاذ القرارات الاستراتيجية، خصوصاً خلال الأزمات. كما تُوفر تحليلاً شاملاً لدور البيانات في تعزيز المرونة والاستجابات المستترة في سيناريوهات المخاطر.

* Corresponding author. E-mail: redouaneali.zineb@univ-constantine2.dz

Doi:

Introduction :

Today, data plays an important role in organizations due to the explosion in the amount of data and the different sources and of its types. Data has become one of the most important assets that organizations possess and plays a critical role in decision making, evaluating organizations and developing plans and strategies. In the financial sector, data is a vital component of risk management, which focuses on the need to ensure its quality in order to achieve the strategic and operational objectives of financial institutions, and the importance of data quality in banks is reflected in global legislation and regulations, particularly those of the Basel Committee on Banking Supervision (BCBS), in particular the BCBS239 document issued by the Basel Committee following the global financial crisis of 2007-2008. The principles of BCBS239 aim to enhance the practice of risk data aggregation and reporting with accuracy and timeliness. This research provides an analytical study of the efforts made by Global Systemically Important Banks (G-SIBs) to comply with the principles of BCBS239, highlighting the critical importance of data quality in compliance, which encompasses data accuracy, reliability, completeness, timeliness, and adaptability.

This study addresses the data quality challenges faced by banks, to tackle these challenges, the research offers recommendations for enhancing data quality, including the improvement of data governance frameworks, the strengthening of data management, and the adoption of modern information technology, by implementing these measures, banks can make accurate decisions, achieve a competitive advantage, gain customer trust, and avoid financial penalties associated with non-compliance with regulations.

I- Data quality

Data quality plays a fundamental role in supporting effective decision-making and achieving organizational objectives. In the context of digital transformation, ensuring data quality has become an imperative to enhance performance, build trust, and achieve competitive excellence. In this section, we will explore the definition and key dimensions of data quality, as well as discuss the data quality management.

1- Definition of data quality

In general, we are attempting to address the historical emergence of the term "data quality," which was first used by statisticians in the late 1960s. At that time, data repetitions in statistical tables were a problem, and researchers were looking for a mathematical theory to identify and prevent these repetitions.

Early in the 1980s, management researchers began to focus on controlling data production systems to identify and fix issues with data quality.

The early 1990s saw its formal and public appearance in the field of computer science, after the rising electronic revolution, the emergence of electronic data, and the development of methods for enhancing and assessing the quality of electronic data kept in antiquated databases, warehouses, and systems.¹

One of the first to define the term data quality is the comprehensive data quality management group from the Massachusetts Institute of Technology, Professor Richard Y. Wang in 1996, which defined data quality as "fitness for use."²

Data is considered to be of high quality if it meets the user's intended requirements. The study of data quality has been based on various dimensions, some of which we will discuss later.

The 1996 Wang also suggested that raw materials could be manufactured and processed as consumer products, as well as data taking the same stages from raw data to processing, improvement, assessment and measurement to consumable data or so-called information."³

It is also defined by J.M.Juran: "as appropriate for data to serve a particular purpose or context in decision-making processes, planning as perceived by users." J.M.Juran is based on five principles: who uses data, how data are used, whether there is a risk to human safety, and what economic resources producers and users of data, what characteristics users take into account when determining the appropriateness of use, use."⁴

Another definition of data quality is: "the distance between the data representations provided by an information system and the same data in the real world."⁵

This means that data quality largely depends on how well the data within an organization, as managed through its information system, matches the data obtained from the real world.

It can be stated that data of high quality is data that fulfills its intended purpose, while it lacks quality if it does not meet the required criteria.

2- Data quality dimensions:

Data quality is usually described as a multi-dimensional concept that is each dimension part of data quality, with data quality dimensions defined by Strong 1996: "as a set of data quality features that represent one aspect or part of data quality."

Before addressing some of the dimensions of data quality, we summarize the approaches based on the diversity of dimensions: the intuitive approach, the theoretical approach and the empirical approach.

- a- **The intuitive approach** is used to determine the quality of data through intuition or so-called experience and is often defines the accuracy of the data, and what characterizes it is that each study can choose the features most relevant to the objectives of the study.
- b- **Theoretical Approach:** This approach derives data quality attributes based on the absence of data defects. Some definitions of data quality use the concept of zero defects, employing Crosby's Absolute Quality Management, which requires data quality to align with its requirements and avoid errors as a means to achieve defect-free data. ⁶One of the advantages of the theoretical approach is its ability to provide a comprehensive set of data quality attributes that are essential for data management.
- c- **Empirical Approach:** The empirical approach to data quality analysis focuses on the users' perspective, asserting that data quality is determined by how the data is utilized by consumers. According to this approach,⁷data quality encompasses several key dimensions by Wang and Strong (1996): ⁸
 - Intrinsic quality: This means that the data must be accurate and reliable;
 - Contextual quality: The data should be relevant and appropriate to the specific temporal and spatial context;
 - Representational quality: This emphasizes the importance of presenting the data in a form that is understandable and easily analyzable;
 - Accessibility quality: This implies that the data should be easily accessible and comprehensible to consumers.

From these approaches, it is evident that there is no general consensus on the dimensions of data quality, with the number of dimensions approaching over 200. It is not feasible to rely on a single dimension to measure data quality. Therefore, we consider the most commonly used and agreed-upon dimensions among researchers.

1-2 Accuracy: Accuracy is the most crucial dimension of data quality. It is defined as the correctness and fidelity of data representation. Accuracy represents the degree of alignment between the data and the real world, reflecting the extent to which data is free from errors. The more accurate the data, the more reliable it is.⁹

It can be measured as follows:

$$\text{Free of error rating} = 1 - \frac{\text{Number of data units in error}}{\text{Total number of data units}}$$

2-2 Timeliness: Timeliness refers to the extent to which data maintains attributes that define its proper duration within a given context. Its primary objective is to ensure that the data remains up-to-date and avoids becoming obsolete.¹⁰

It is measured as follows:

$$\text{Timeliness} = \text{Data delivery time} - \text{Occurrence time or (Data provided)}$$

3-2 Consistency: Consistency can be viewed from several perspectives. For example, Cod's referential integrity constraint represents the consistency of redundant data within a single table or across multiple tables. The second type of consistency relates to the alignment between two related data elements. The third type focuses on standardizing the same data element when used in different tables.¹¹

It is mathematically represented as follows:

$$\text{Consistency rating} = 1 - \left(\frac{\text{Number of instances violating specific consistency type}}{\text{Total number of consistency checks performed}} \right)^{12}$$

4-2 Completeness: Completeness is generally described as the breadth, depth, and scope of the information contained within the data. It represents the extent to which the required data is present, created by Blake and Mangiameli in 2011, which consist for a relation R , let TR be the number of tuples in which have at least one 'NULL'-value and let NR be the total number of tuples in R . Then, the completeness of R is defined as follows :¹³

$$\text{Completeness} = 1 - TR / NR = NR - TR / NR$$

5-2 Integrity: Integrity refers to the state of completeness and cohesion or uniformity. It is the degree to which data conforms to relational rules intended to ensure a comprehensive, consistent, and accurate representation of the data.¹⁴

It is measured as follows:

$$\text{Integrity \%} = (\text{Number of records that have valid relationships}) \times 100 / (\text{Number of records with missing linkages} + \text{Number of records that have valid relationships} + \text{Number of records that have invalid relationships})^{15}$$

3- Data quality management:

Data and information are the means from which decisions, analyses and planning are concluded, and data quality management is a series of people management processes, policies and technology aimed at improving the assessment and assurance of data quality, and the primary objective of institutions' adherence to data quality management is to ensure operational efficiency and compliance with legal and regulatory requirements, customer satisfaction, reputation, competitive advantage and sound decision-making.

The mature data quality management programme, which takes into account data quality issues and attempts to reduce bad data, and here reference must be made to The 1-10-100 rule developed by George Labovitz and Yu Sang Chang in 1992, which is considered to cost the verification of the record after the entry of the data \$1, \$10 to clean up and eliminate repetition after entry, but \$100 in productivity or potential revenue if no action is taken.¹⁶ It applies quality management techniques to data, ensuring its suitability for use and alignment with the needs of data consumers.¹⁷

To achieve high-quality data, the Six Sigma methodology can be implemented, it developed by Motorola in 1980¹⁸, Six Sigma is a quality management program designed to fundamentally improve operational performance by designing and monitoring business activities in ways that reduce resource waste. It facilitates rapid improvements in cost, quality, process speed, safety, customer satisfaction, investment return, and environmental impact. The term "Sigma" refers to standard deviation, emphasizing the goal of reducing variation. Six Sigma was initially developed as a statistical methodology to enhance the DMAIC process, with the primary objective of achieving processes and products that are nearly defect-free.¹⁹ The DMAIC process consists of five key phases aimed at improving data quality and solving problems in a systematic and sustainable manner, these stages are: definition, measurement, evaluation, improvement and control. Some references adopt the DAIC method (definition, assessment, improvement, control) which is the measurement phase within the evaluation phase.

The phases of the DMAIC methodology are as follows:

Definition phase: The primary and important phase of data quality management is based on the identification of organizational requirements and the integration of data and key points into data flows, the development of a road map for data access and data ranking, the data collected must be integrated into data governance tools, and the data governance is comprised of policies, technologies, people responsible for data management, determining the basis for decision-making, how to comply with laws and regulations.²⁰

Measurement phase: To assess the data quality level against the set objectives, data analysts measure the dimensions identified in the first phase or those closely related to the organization's requirements, the measurement can be either quantitative or qualitative, we have previously discussed the dimensions of data quality and how they are measured.²¹

Analysis phase: This stage aims to analyze the high-level solutions identified during the measurement stage and to create and design detailed processes for each solution,²² in this phase, data profiling is conducted, which involves a set of methods to examine and analyze a range of databases and metadata to identify potential root causes, this process typically relies on statistical analysis to determine the number of missing values, distinct values, and data types.²³ Identified causes can be visualized using cause-and-effect diagrams, also known as Ishikawa diagrams or fishbone diagrams, this process consolidates the causes, aiding in the identification and classification of the impact of root causes on the problem.²⁴

Improvement phase: This phase focuses on finding solutions to eliminate the root causes of the problem by working with IT teams to automate and streamline the improvement process.²⁵

Control and Monitoring phase: It involves data monitoring, which entails continuously and regularly examining databases and data warehouses against established data quality standards, the results are displayed on a data quality dashboard, which issues alerts if there are issues or breaches in data quality rules, this allows for the rapid detection and correction of problems before poor data accumulates. Stakeholders and IT teams are responsible for defining data

quality management requirements. Monitoring can be automated through specialized software, manually via SQL and Excel, or through a combination of both methods.²⁶

II-BCBS239:

In light of banks' inadequate response to risks due to deficiencies in information technology and weak data structures factors that significantly contributed to the global financial crisis of 2007/2008 BCBS239 focuses on banks' ability to manage risks. To prevent such crises in the future, the Basel Committee on Banking Supervision (BCBS) issued the BCBS239 document in January 2013. This document emphasizes and strengthens the capability to aggregate risk data and practice risk reporting. BCBS239 was designed to reinforce the second pillar (supervisory review) of Basel II and III, aiming to enhance banks' ability to identify and manage risks. It addresses the necessary improvements to avoid future crises.²⁷

1- The principles of BCBS239 :

The BCBS239 document consists of principles and standards directly targeting Global Systemically Important Banks (G-SIBs), while other banks are also affected and can assess themselves, it includes 14 principles divided into four groups. The aim is to provide a roadmap for banks and mandate compliance with 11 of these principles, while the remaining three principles pertain to supervisory review, tools, and collaboration. Based on the BCBS239 document, the 14 principles are summarized as follows:

Section One: Overarching Governance and Infrastructure:

Banks are required to establish robust data governance, risk data infrastructure, and information technology systems, which together form the cornerstone of their capacity to manage and mitigate risks. The stronger and more resilient a bank's data structures particularly for risk data and the higher the quality of its IT infrastructure, the better positioned it will be to confront and withstand risk. Additionally, the presence of strong governance capable of effective management and oversight enhances a bank's ability to face challenges. The first group of principles is founded on three key pillars:

Table N°1: The principles of section one

Principle 1 : Governance	The bank's capability to consolidate risk data and its risk reporting practices must be governed by a robust framework that aligns with the principles and guidelines established by the Basel Committee.
Principle 2 : Data architecture and IT infrastructure	A bank must build and sustain a robust data architecture and IT infrastructure to support risk data aggregation and reporting, ensuring efficiency under all conditions and compliance with established principles.

Source : BCBS239, Principles for effective risk data aggregation and risk reporting, Basel committee on banking supervision, January 2013, P13, 14.

Section Two: Risk Data Aggregation Capabilities:

This refers to a bank's ability to collect risk data, the data plays a fundamental role, as decisions and risk predictions cannot be made without it. Banks are obligated to prioritize high data quality when gathering their data, which is essential for effective risk management. Higher-quality data leads to more accurate risk assessments and the more reliable the decision-making process, enabling forecasters to provide insights and predictions about future risks or profitability.

To manage risks effectively, BCBS239 emphasizes that risk data must be both accurate and timely, which is achieved through high data quality. This group consists of four principles that represent dimensions of data quality. BCBS239 places significant importance on data quality due to its crucial role in risk management within banks. These principles are:

Table N°2: The principles of section two

Principle 3 : Accuracy and Integrity	A bank must ensure accurate and reliable risk data for reporting under all conditions, with an automated data aggregation process to minimize errors.
Principle 4 : Completeness	A bank must ensure the ability to gather and integrate critical risk data from all areas of its organization, categorizing it by business line, legal entity, asset type, industry, region, and other pertinent factors. This classification allows for effective identification, reporting, and analysis of risk exposures, concentrations, and emerging risks.
Principle 5 : Timeliness	A bank must generate timely, accurate, and complete aggregated risk data, ensuring it aligns with principles of integrity and adaptability. The timing of this data depends on the nature and volatility of the measured risk, its significance to the bank's overall risk profile, and the bank's specific reporting requirements, both under normal and stressed conditions.
Principle 6 : Adaptability	A bank must be capable of generating aggregated risk data to meet various on-demand reporting requirements, including those triggered by stress or crisis situations, evolving internal needs, and supervisory queries.

Source : BCBS239, Principles for effective risk data aggregation and risk reporting, Basel committee on banking supervision, January 2013, PP 15-17.

Section Three: Risk Reporting Practices:

For the board of directors and senior management to make informed decisions, accurate and timely reports are essential. These reports must be based on high-quality risk data that is clear and comprehensive, covering all aspects of risk. They provide decision-makers with a roadmap for addressing risks. This group includes five principles concerning the necessary attributes of risk reports for banks, ensuring credibility and reliability to support sound decision-making. These principles are as follows:

Table N°3: The principles of section three

Principle 7 : Accuracy	Risk management reports must present aggregated risk data with accuracy and precision, ensuring a clear and exact representation of risks. These reports should undergo thorough reconciliation and validation processes to maintain their reliability.
Principle 8 : Comprehensiveness	Risk management reports should comprehensively address all significant risk areas, with their depth and scope aligned to the bank's size, operational complexity, risk profile, and the needs of their recipients.
Principle 9 : Clarity and usefulness	Risk management reports should convey information in a clear and concise manner, balancing risk data, analysis and qualitative insights. They should be comprehensive enough to support informed decision making and tailored to the specific needs of their recipients.
Principle 10 : Frequency	The board and senior management should determine the frequency of risk management reports based on the needs of the recipients, the nature and volatility of the risks, and the role of the reports in supporting informed decision-making. The frequency of reporting should be adjusted and increased during periods of stress or crisis to ensure effective risk management.
Principle 11 :	Risk management reports should be distributed to the appropriate parties while maintaining confidentiality.

Source : BCBS239, Principles for effective risk data aggregation and risk reporting, Basel committee on banking supervision, January 2013, PP 18-21.

Section Four: Supervisory Review, Tools, and Collaboration:

This group focuses on the role of supervisors in tracking, monitoring, overseeing, providing assistance, and incentivizing banks to comply with the abovementioned 11 principles. Supervisors are responsible for guiding banks on the best practices for comprehensive compliance and ensuring effective implementation of these principles. This group consists of four principles, which are:

Table N°4: The principles of section four

Principle 12 :	Supervisors should conduct periodic assessments to evaluate a bank's compliance with the eleven principles.
Principle 13 :	Supervisors should have the necessary tools and resources to enforce timely corrective action for deficiencies in a bank's risk data aggregation and reporting. They should have access to a range of mechanisms, including those under Pillar 2, to ensure compliance and effectiveness.
Principle 14 :	Supervisors should cooperate with their counterparts in other jurisdictions to monitor the application of the principles and take corrective action where necessary.

Source : BCBS239, Principles for effective risk data aggregation and risk reporting, Basel committee on banking supervision, January 2013, P21, 22.

The Basel Committee set January 2016 as the target date for full compliance with BCBS239 principles for G-SIBs. It was evident that banks would face challenges in achieving full compliance during the initial implementation phase. However, after 10 years since its inception and 7 years past the deadline, reports on the status of G-Sib's adherence to BCBS239 reveal that only two banks have achieved full compliance with the principles, with no single principle reaching full compliance across all banks.

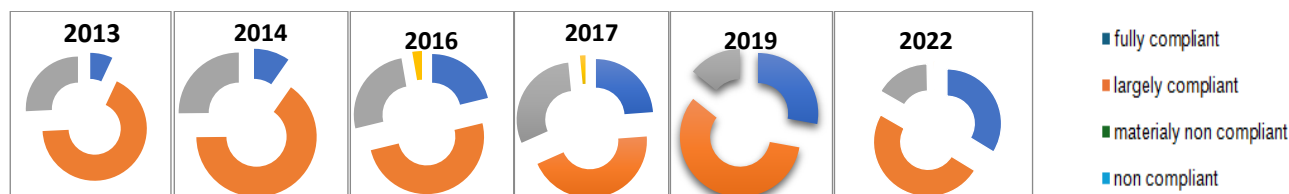
2- Objectives of BCBS239

Based on the lessons learned from the global financial crisis, BCBS239 was established as a fundamental reference for banks to address and prevent future crises; the document aims to achieve the following objectives:

- **Improve Risk Data Quality:** Enhance the quality of risk data, which plays a crucial role in making well-informed decisions;²⁸
- **Analyze, Monitor, and Forecast:** Analyze, monitor, and anticipate current or future challenges, whether they pertain to profitability, crises, or stress situations, and provide insights and strategies for effective decision-making and future planning;²⁹
- **Strengthen Governance:** Guide banks in strengthening and enhancing their governance frameworks, emphasizing the importance of robust management structures that facilitate effective data flow from the operational level to senior management;³⁰
- **Ensure Robust Infrastructure:** Ensure that banks have strong data infrastructure and technology capable of handling and analyzing large volumes of data effectively;
- **Encourage Comprehensive Risk Data Collection:** Urge banks, especially large and global institutions with multiple branches, to gather risk data comprehensively and avoid data silos;
- **Promote Automation in Data Collection:** Advise banks to reduce manual data collection processes and shift towards automated systems to minimize errors and achieve higher data quality;
- **Enhance Risk Management:** Strengthen risk management practices and decision-making processes within banks;³¹
- **Improve Reporting Structures:** Emphasize the importance of robust information reporting frameworks and ensure that risk reports effectively convey accurate information to decision-makers;³²
- **Prioritize Data Quality:** Reinforce the principle that accurate data leads to correct decisions, while incorrect data leads to faulty decisions, underscoring the need for high-quality risk data and its application;
- **Ensure Compliance:** Encourage banks to comply with the principles and provide adequate time for compliance to foster financial stability;³³
- **Advance Strategic Planning:** Improve the quality of strategic planning and the bank's ability to manage risks associated with new products and services.

3- Compliance of G-SIB with BCBS239 Principles

We analyzed data and reports based on BIS to assess the compliance of G-SIBs with BCBS239 principles, as illustrated in the following figure:

Figure N°1: The compliance of G-SIBs with BCBS239 principles

Source: The diagram was created based on the reports issued by the BIS reports

Thirty banks were self-assessed by supervisors, with each principle evaluated individually, all assessments were compiled into BIS reports in the first assessment of banks in 2013 a positive response was observed, notably with no bank being non-compliant with the principles, it was found that 67% of the banks were largely compliant. However, only 7% of the banks evaluated themselves as fully compliant with the principles, indicating the need for intensified efforts to achieve full compliance.

The same trend was observed in 2014, with slight variations, the percentage of full compliance increased to 10%, while the percentage of largely compliant banks slightly decreased to 65%, this is considered relatively good progress.

For the years 2016 and 2017, it is noted that the percentage of non-compliant appeared after being absent in the previous reports. This is due to the challenges banks faced with some principles, which will be detailed later. On a positive note, the percentage of full compliance increased to 21% in 2016 and 24% in 2017, which is a good indicator for G-SIB banks.

In the 2019 report, banks improved their performance, raising the percentage of full compliance to 28%, reducing material non-compliance to 14%, eliminating non-compliant, and increasing the percentage of largely compliance to 58%.

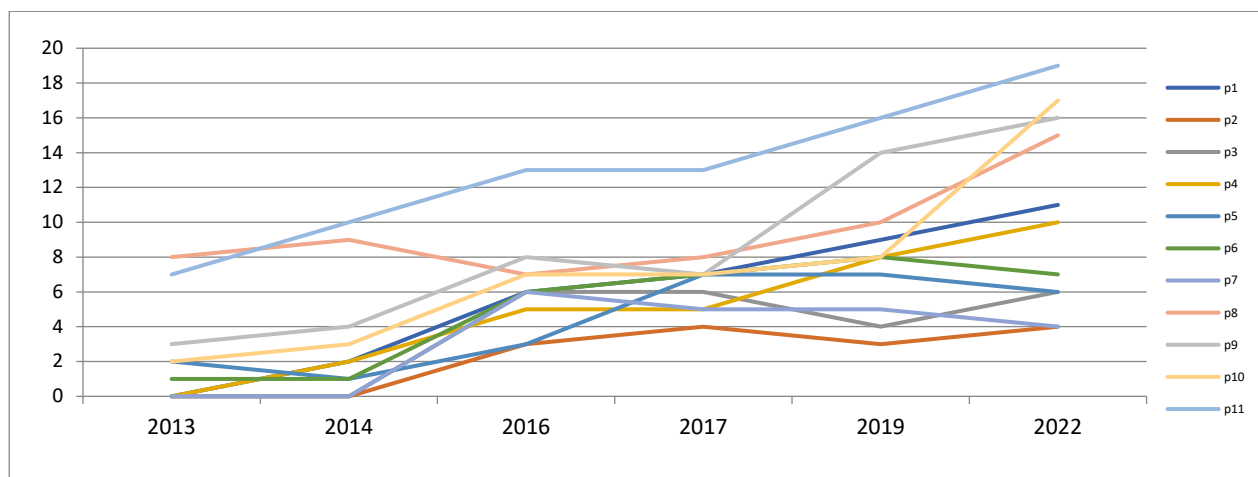
After the COVID-19 pandemic and global events, in the year 2022, there is a noticeable improvement in bank compliance; this indicates the banks' ability to avoid risks and crises, and their commitment to global financial regulations, including BCBS239.

4- Comprehensive Assessment of BCBS 239 fully compliant: Insights into the Performance of G-SIB

As previously discussed, the BCBS239 principles have a logical and coherent sequence that is interlinked, meaning that no principle can be prioritized over another. Principles 1 (Governance) and 2 (Data Infrastructure) are foundational for the implementation of the remaining principles. This makes it crucial for banks to diligently strive for compliance with these two principles to ensure that the other principles can be effectively adhered to.³⁴

Based on the six BIS reports on the application of BCBS239 principles by G-SIBs, we aim to provide a comprehensive overview of how banks conduct self-assessment in their adherence to these principles.

In the following figure, we present a graphical representation of the 11 principles from the six reports regarding banks' compliance with BCBS239. We will discuss each principle individually.

Figure N° 2: Fully compliant of the principles

Source: The diagram was created based on the reports issued by the BIS reports.

Principle 1: Governance

The Basel Committee, particularly through the BCBS239 document, emphasizes the necessity of adhering to a robust governance framework and involving the board and the senior management of risk data, as it is a fundamental component of risk management.

A bank is considered compliant with this principle when it has established effective policies and procedures for managing risk data, clearly defining ownership and responsibilities to ensure accountability for data quality. However, the six reports indicate a significant weakness in banks' adherence to this principle, categorizing it as a level two non-compliance. Since the beginning of the reports, no bank achieved full compliance with the governance principle. Nevertheless, as banks recognized the critical importance of this principle, they made substantial efforts, which are evident in the positive development of the number of banks achieving full compliance, reaching 11 banks by 2022.

Principle 2: Data architecture and IT infrastructure

Principle 2, focusing on data infrastructure, is a critical element for banks, not only for complying with the principles but also for enhancing their competitive edge, especially in light of rapid advancements in information technology. A robust data infrastructure is essential, supported by advanced IT systems that can quickly adapt to new developments. This includes automated data entry and report generation, which help minimize errors and ensure data quality.

However, this principle is the least complied with compared to others. Despite banks' efforts to enhance their IT infrastructure, it has not met the Basel Committee's expectations, particularly regarding preparedness for crises and stress situations. No bank achieved full compliance in the 2013 and 2014 reports. In subsequent reports, only between 3 and 4 banks achieved full compliance, making it the least adhered to principle.

Principle 3: Accuracy and integrity

Principle 3 requires banks to collect risk data accurately and integration from their sources, ensuring it is governed by stringent data quality controls. Data accuracy should always be measured and monitored, with plans in place to address poor data quality. A key factor in achieving data accuracy is having a well-structured data framework and advanced IT systems that enable automated data entry to avoid errors. Additionally, a sophisticated data warehouse is necessary to facilitate the organization, classification, and segmentation of data according to its usage and need.

It is noteworthy that no bank achieved full compliance for the periods of 2013 and 2014. However, the number of banks meeting full compliance increased to 6, remaining stable, except for 2019 when only 4 banks fully complied with this principle.

Principle 4: Completeness

Principle 4 emphasizes that banks must strive to provide all relevant data pertaining to their operations, particularly concerning risk, to take necessary actions. Data with completeness provides a comprehensive view of the bank's situation, aiding in the avoidance of potential risks and crises. This principle measures the extent to which data is complete and comprehensive.

From the graphical representation, it is observed that no bank fully complied with this principle in 2013. However, banks recognized its importance and, subsequently, the number of fully compliant banks increased to 2 in the following year and eventually rose to 10 banks in subsequent years.

Principle 5: Timeliness

Principle 5 addresses the banks' ability to provide risk data in a timely manner, especially during critical times such as crises and periods of stress. This data must always be accurate, reliable, and complete across all banking operations. The principle heavily relies on the speed of data entry and flow, with automated data entry significantly enhancing the prompt availability of information.

No bank fully complied with this principle in 2013. Compliance improved to one bank in 2014, and increased to a maximum of 7 banks between 2017 and 2019. However, the number of fully compliant banks decreased to 6 by 2022.

Principle 6: Adaptability

Principle 6 focuses on the banks' ability to swiftly adapt and integrate new developments, particularly in addressing critical situations and crises. It is crucial to avoid relying on manual solutions to meet specialized requirements effectively.

The graphical representation indicates that only one bank was fully compliant with this principle in 2013 and 2014. However, banks improved their compliance, reaching a maximum of 8 fully compliant banks in 2019, before decreasing to 7 banks by 2022.

Principle 7: Accuracy

Principle 7 emphasizes that reports must be accurate and correctly reported, which directly depends on the quality of risk data. It also highlights the importance of minimizing reliance on manual processes for preparing risk reports.

No banks were fully compliant with this principle in 2013 and 2014. Compliance improved to a maximum of 6 banks in 2016 but subsequently decreased to just 4 banks by 2022. This indicates a concerning status regarding reporting practices within banks.

Principle 8: Comprehensiveness

Principle 8 requires that detailed information on all risks surrounding banks be provided to facilitate necessary actions and scenario analysis during stress testing.

The graphical representation shows that the number of banks fully compliant with this principle was reasonably high in the initial period, reaching 8 banks. However, compliance dropped to 7 banks in 2016, before rising to 15 banks in 2022.

Principle 9: Clarity

Principle 9 mandates that risk reports must be clear and easily interpretable to facilitate analysis and provide a clear view of the information needed by the board and senior management for decision-making.

It is observed that there has been a gradual improvement, with the number of banks fully compliant increasing from 3 in the initial period to 16 by 2022.

Principle 10: Frequency

Principle 10 emphasizes that risk reports directly contribute to risk management by supporting decision-making, particularly during critical times. It is noted that banks have improved their compliance significantly, with the number of banks fully compliant increasing from 2 in 2013 to 17 by 2022.

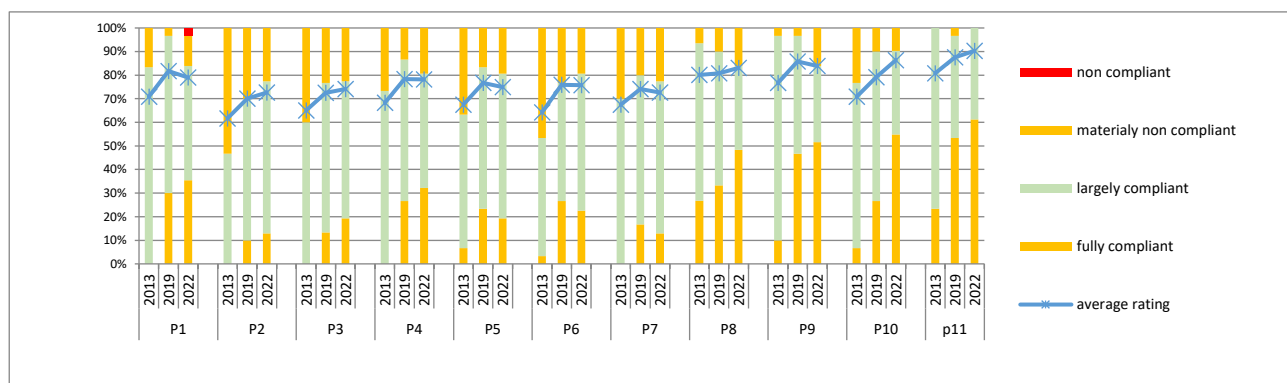
Principle 11: Distribution

Principle 11 focuses on the distribution of risk reports to relevant stakeholders while maintaining complete confidentiality. It is observed that banks have effectively adhered to this principle, with compliance increasing from 7 banks in 2013 to 19 banks in 2022. Consequently, this principle is considered the most compliant.

5-Challenges for G-SIBs in Complying with BCBS239 Principles during Crises

We aimed to briefly examine the ability of banks to meet compliance with BCBS239 principles during crises. The opportunity to test this arose during the global pandemic and the stress events related to the war in Ukraine, as many G-SIBs are based in countries connected to the conflict. The following graph focuses on comparing three key compliance periods: the initial compliance period in 2013, the pre-crisis period in 2019, and the post-crisis period as reported in 2022.

Figure N°3: Challenges for G-SIBs in Complying with BCBS239 Principles during Crises



Source: The diagram was created based on the reports issued by the BIS reports

When considering the average compliance rates, we observe notable changes between the periods of 2013, 2019, and 2022. There is a decrease in compliance for principles 1, 5, 7, and 9, with a slight decline in principle 6. Principles 8 remain relatively stable, while principles 2 and 3 show a slight increase.

- **Non-compliance:** It is noted that no bank achieved non compliance across the three periods for all principles except for all Principles, with the exception of Principle 1 in 2022, which had a non-compliance rate of 3.22%.
- **Material Non-Compliance:** There was a decrease between 2013 and 2019, followed by an increase in 2022 for the following principles: 1,4, 5, 7, and 8. In contrast, there was a decline across all three periods for the following principles: 2, 3, 6, 9, and 10. Principle 11 saw a transition from no bank material non-compliant in 2013 to an increase in 2019, only to revert to no bank material non-compliant again in 2022.
- **Largely Compliant:** There was a decline in all principles except for a rise in Principle 3 in 2019, and an increase in Principle 7 in 2022 compared to 2019.
- **Fully Compliant:** A general increase is observed across all principles to a small degree. However, Principles 4, 5, 6, and 7 Principles 7, 6, 5, and 4 declined in 2022 compared to 2019

Overall, it is noteworthy that the principles most affected by the global crises were 1, 5, 7, and 9. In general, the COVID-19 pandemic and other crises tested the effectiveness of the BCBS 239 principles in ensuring the quality of risk data management in banks and financial institutions. As most banks transitioned to remote work during the pandemic, a significant challenge in implementing full compliance with BCBS 239 principles. That encourages banks to strengthen their information systems and technological infrastructure to aggregate and integrate high-quality data from a variety of sources.³⁵

III-Data Quality in BCBS239

1. The importance of data quality in BCBS 239

Data quality holds significant importance in the financial sector due to the regulations and global standards that emphasize the need for continuous improvement in data management programs. Consequently, banks and financial institutions must collect, measure, assess, enhance, and monitor high-quality data. This is in line with the compliance requirements of financial regulations such as Basel II and Basel III, which underscore the necessity for complete transparency and traceability of data to identify potential risks that banks may face. The BCBS 239 document emphasizes this point, with its primary goal being to enhance banks' capabilities in aggregating risk data (such as interest rate risk, credit risk, operational risk, market risk, liquidity risk, etc.) and reporting it effectively. This, in turn, strengthens banks' ability to manage risks and make informed decisions.

The importance of data quality has been highlighted due to the necessity of enforcing strict controls on data quality as one of the requirements of BCBS 239.³⁶ Managing data quality has become increasingly challenging as data has become a crucial resource in contemporary times. The role of data analysts and data stewards has become as significant as that of accountants and financial officers. Data analysts often spend considerable time gathering data from multiple sources, attempting to improve inconsistent, duplicate, or incomplete data, and preparing it accurately and reliably in a timely manner for decision-making. By addressing these key challenges, banks and financial service providers are better equipped to manage risks and make informed decisions.

To ensure successful compliance with BCBS 239, banks must prioritize maintaining high-quality data, the direct significance of data quality in BCBS 239 is clearly highlighted through Principles : 3 (Accuracy and Integrity), 4 (Completeness), 5 (Timeliness), and 6 (Adaptability). These principles represent the essential dimensions of data quality, and they are interrelated, with each influencing and being influenced by the others within the framework.

2. Data Quality Challenges in BCBS239

Poor data quality is one of the biggest obstacles banks face, significantly impacting decision-making and leading to loss of credibility, competitive advantage, and customer trust. Therefore, banks must first address the consequences to remain competitive in the market and, second, comply with BCBS239 principles. Data quality plays a crucial role in reporting, as poor data can have negative effects on decision-making, particularly in risk management. It may lead to regulatory reporting errors. Data quality issues can arise from incomplete data, low accuracy, or inconsistency. Additionally, data duplication, where multiple copies of the same data exist, can cause discrepancies, including measurement and reporting errors.

Data quality challenges within G-SIBs lead to a decrease in their compliance capabilities. Among these challenges are:

- Most banks encounter difficulties in improving data quality, with many issues attributed to the lack of controls in front-office operations. Inefficiencies in data structures and data infrastructure are often present in outdated systems that still rely on manual data entry processes. This leads to operational shortcomings and makes enhancing data quality more challenging;³⁷
- The multitude and variety of data sources used to aggregate risk data necessitate complex operational processes to ensure its quality and suitability for use. Typically, identifying the root cause in the underlying data is challenging;³⁸
- One of the primary causes of data quality issues is manual data entry, which stems from the outdated IT infrastructure of legacy systems. This process requires significant effort, substantial financial costs, and several years to develop and upgrade IT systems within banks;
- The lack of a robust data quality framework (data quality rules, controls, and processes), inadequate data management (data owner and steward responsibilities), and the complexity of data dictionaries are among the reasons why banks struggle to meet the requirements of the BCBS239 framework;³⁹
- Monitoring of risk data is inadequate and there is no methodology for compiling risk data from documents available to the board and senior management;
- The process of collecting data and preparing risk reports in most cases is done manually, which affects the timing, especially in stressful situations;⁴⁰
- Large banks, especially global ones with legal jurisdictions, are often affected by data silos, where data is isolated in separate systems or departments with insufficient interaction or integration. This isolation leads to issues with data access, data redundancy, and difficulties in achieving a comprehensive and accurate view. Banks also face challenges in standardizing data definitions, handling incomplete datasets, and lacking uniform

data management practices. All these factors contribute to the complexity of analysis and decision-making processes, highlighting the need for effective data integration and quality improvement strategies;⁴¹

- There is a need to design and maintain data lakes that are scalable and adaptable. Managing data lakes requires significant effort to ensure data integrity and accuracy, especially when dealing with multiple and diverse data sources;
- The unification of systems and the provision of purpose-driven data are hindered by the insufficient data lineage throughout its lifecycle.⁴²

3. Strategies for Enhancing Data Quality to Achieve Principle Compliance

Most banks focus on enhancing data management and data quality to meet the requirements for compliance with BCBS 239. We have formulated several strategies based on BIS reports, outlined in the following key points:

To achieve full compliance with the principles, banks should establish clear, long-term roadmaps rather than relying on tactical solutions. These roadmaps should include the following:

- Researching and identifying potential compliance gaps and developing solutions to address them;
- Advanced technology helps banks address ongoing challenges such as automating documentation, reducing manual interventions, and automating data flows;
- Designing and maintaining data lakes to enhance scalability and adaptability is crucial to meeting the growing and diverse needs of stakeholders;⁴³
- It is preferable to link data quality management processes with the business operations, risk, finance, and information technology and to ensure the verification of data quality, including the correction of existing errors;
- Enhance data quality, and improve data discovery and characterization to identify and measure data directly used in risk management and regulatory compliance;
- Develop tools for data quality monitoring and proactive oversight to detect critical data errors, especially those related to compliance and alert users accordingly;
- Precisely analyze data ratios from start to finish and prepare the necessary reports to meet regulatory audit requests. This helps avoid the burdens associated with manual documentation of data collection, transformation, and cleansing, filtration, and calculation processes;
- Develop comprehensive roadmaps that focus on long-term initiatives, as opposed to relying on tactical solutions, to address data quality and aggregation challenges.⁴⁴

Conclusion

Data quality plays a fundamental role in enhancing banks' ability to manage risks and comply with regulations, particularly the BCBS239 principles, which are clearly articulated through the dimensions of accuracy, reliability, adaptability, timeliness, and completeness. The analytical study of BIS reports on the efforts of G-SIBs to comply with BCBS239 principles highlights the challenges banks face in achieving and improving data quality. The study aimed to elucidate the strengths and weaknesses of G-SIBs in meeting these compliance requirements and to propose solutions for achieving full compliance. The conclusions drawn from this analysis are as follows:

- The importance of data quality in compliance with principles, particularly since several BCBS239 principles are fundamental dimensions of data quality;
- One of the main challenges banks faced in improving their data quality included issues related to governance, data structuring, and information technology;
- Implementing data quality helps banks identify potential risks and handle them more effectively by making timely and accurate decisions;
- Data quality aids in meeting regulatory requirements accurately and efficiently, thus preventing financial penalties imposed by regulatory in cases of non-compliance;
- Data quality contributes to reducing errors throughout the data lifecycle, thereby increasing operational efficiency;
- Banks have sought to enhance senior management accountability to improve data management and raise awareness about data quality. Additionally, the accuracy, reliability, and timely availability of reports bolster stakeholder confidence.

The recommendations of this study include the following:

- Establishing a robust governance framework and accountability for data ownership to ensure data quality.
- Utilizing modern information technology and enhancing data infrastructure;

- Employing advanced data analysis tools, such as machine learning, to facilitate the detection of data quality gaps and address them with minimal effort and in a timely manner;
- Enhancing data quality management throughout the data lifecycle, from data collection to the production of information necessary for decision-making;
- Data architecture, quality controls, and governance mechanisms should support banks in risk management and add business value, enabling banks to achieve high standards of data quality in risk management and regulatory compliance;
- Coordinated efforts can be implemented to automate processes, numerize operations, and develop advanced analytical capabilities to strengthen banking activities;
- Banks must develop their capabilities in data aggregation and reporting with accuracy and timeliness, especially when facing crises and risks;
- It is essential to create a system that effectively evaluates, identifies, and manages data-related risks to avoid losses resulting from decisions made based on poor-quality data,
- Proactive monitoring and oversight of data quality enable a swift and effective response to identify root causes, avoid errors, and promptly address issues requiring immediate intervention,
- Artificial intelligence enables the enhancement of data quality standards by examine of the data being collected and utilized, in alignment with the BCBS 239 standards, which define data quality as fundamental requirements.⁴⁵

By implementing these recommendations, banks can enhance their data quality and improve their ability to comply with regulations and manage risks effectively. Adhering to regulations and legislation, coupled with making well-informed decisions, boosts customer confidence, strengthens competitive advantage, and maintains financial stability in the face of crises.

Referrals and References:

- 1 Carlo Batini, Monica scannapieca, Data quality :concepts, methodologies and techniques, Springer, Berlin, 2006,P4.
- 2 D.B.Shanmugam, J.Dhilipan, A.Vignesh,et al, Challenges in data quality and complexity of managing data quality assessment in big data, international journal of recent technology and engineering(IJRTE),volume 9, September 2020, DOI:10.35940/ijrte.c5643.099320, P590.
- 3 Richard Wang, Yair Wand, Anchoring data quality dimensions in ontological foundations, Communication of ACM, volume 39, November 1996, <https://doi.org/10.1145/240455.240479>, P86.
- 4 Sadia vancauwenbergh,Data quality management, Scientometrics recent advances, Volume 1, Jan 2019, DOI:10.5772/INTECHOPEN.86819, P3.
- 5 Maria del Pilar Angeles, Francisco Garcia Ugalde, A data quality practical approach, international journal on advances in software, Volume 2, 2009, P260
- 6 Sadia vancauwenbergh,Data quality management, Scientometrics recent advances, Volume 1, Jan 2019, DOI:10.5772/INTECHOPEN.86819, P3.
- 7 Jingran Wang, Yi Liu, Peigong Li, et al, Overview of Data Quality: Examining the Dimensions, Antecedents, and Impacts of Data Quality, Journal of the Knowledge Economy (2024), doi.org/10.1007/s13132-022-01096-6,P1165.
- 8 Richard Y. Wang, Diane M. Strong, Beyond accuracy : what data quality means to data consumers, journal of management information systems, Volume 12, 1996, DOI:10.1080/07421222.1996.11518099, P7.
- 9 Lisa Ehrlinger, Elisa Rusz, and others, A survey of data quality measurement and monitoring tools, <https://doi.org/10.48550/arXiv.1907.08138> ,July2019, p7.
- 10 Youdi Gong, Guangzhen Liu, Yunzhi Xue, et al, A survey on dataset quality in machine learning, Information and Software Technology,Elsevier, 162(2023), <https://doi.org/10.1016/j.infsof.2023.107268>, p6.
- 11 David Loshin, The practitioner's guide to data quality improvement,Elsevier, Usa, 2011, p141
- 12 Yang W. Lee, Leo L. Pipino, James D. Funk, and Richard Y. Wang, Journey to Data Quality, The MIT Press, London, England, 2006, <https://doi.org/10.7551/mitpress/4037.001.0001>, P57.
- 13 Berand Heinrich, Diana Hristova, Mathias Klier, et al, Requirements for Data Quality Metrics, ACM Journal of Data and Information Quality, 2018, <https://doi.org/10.1145/3148238> P16.
- 14 Laura Sebastian-Coleman, Measuring Data Quality for Ongoing Improvement, Elsevier,United States of America, 2013,P63.
- 15 Rupa Mahanti, Data quality : dimensions, measurement, strategy, management, and governance, ASQ Quality press, USA, 2018, P315.
- 16 Andrew Jones, Data quality in the age of AI ; Packet publishing, Brimingham ; 2024 ; P19.
- 17 Henny Sri Indriany, Achmad Nizar Hidayanto, Lellyana Juliet Wantania , et al, Data Quality Management Maturity: Case Study National Narcotics Board, IEEE International Conference on Communication, Networks and Satellite (Commnetsat), 2021, P207.

- 18 S. Grace Prasana, P. Subasri, M. Vigneshwaran, The conceptual framework of six sigma, Shodha Prabha (UGC CARE Journal), Vol. 48, Issue 01, No. 6 : 2023,P 2.
- 19 Rajesh Jugulum, Competing with high quality data : Concepts, tools, and techniques for building a successful approach to data quality, John Willey & son, Canada, 2014, P 24.
- 20 Abraham, R., vom Brocke, J., Schneider, J., Data Governance: A conceptual framework, structured review, and research agenda, in: International Journal of Information Management (IJIM), 2019, <https://doi.org/10.1016/j.ijinfomgt.2019.07.008>, P17.
- 21 Sadia vancauwenbergh, Data quality management, Scientometrics recent advances, Volume 1, Jan 2019, DOI:10.5772/INTECHOPEN.86819, P10.
- 22 Mohit Panwar, Application of six-sigma for data quality : improvement in an insurance company, International journal of scientific & technology research, Volume 4, ISSUE 09, September 2015, P4.
- 23 Laura Sebastian-Coleman, Measuring Data Quality for ongoing improvement, Elsevier, USA, 2013, p100.
- 24 Luca Liliana, A new model of Ishikawa diagram for quality assessment, 20th Innovative Manufacturing Engineering and Energy Conference (IManEE 2016), Kozani, Greece , DOI:10.1088/1757-899X/161/1/012099, P2.
- 25 Lea M. Monday, Define, Measure, Analyze, Improve, Control (DMAIC) Methodology as a Roadmap in Quality Improvement, Global Journal on Quality and Safety in Healthcare, Volume 5, Issue 2, April 2022, DOI: 10.36401/JQSH-22-X2, P 45.
- 26 Rupa Mahanti, Data quality: dimensions, measurement, strategy, management, and governance, ASQ Quality press, USA, 2018, P346.
- 27 BCBS239, Principles for effective risk data aggregation and risk reporting, Basel committee on banking supervision, January 2013, P8.
- 28 Report on the Thematic Review on effective risk data aggregation and risk reporting, European Central Bank, 2018, P2.
- 29 Guide on effective risk data aggregation and risk reporting, ECB, July 2023, P11.
- 30 BCBS239, Principles for effective risk data aggregation and risk reporting, Basel committee on banking supervision, January 2013, P10.
- 31 Sean Smith et al, BCBS 239 Compliance: A catalyst for gaining competitive advantage, Deloitte, 2017, P4.
- 32 BCBS239, Principles for effective risk data aggregation and risk reporting, Basel committee on banking supervision, January 2013, P10.
- 33 Jonathan Riva et al, BCBS 239 – Raising the standard, PwC, November 2017, P15.
- 34 BCBS d501, Progress in adopting the Principles for effective risk data aggregation and risk reporting, April 2020, BIS, P3.
- 35 Tuk Yulianto, Tri Ratnawati, Ulfi Pristiana, The Impact of Credit Information Technology, Risk Management Strategy, Compliance, and Restructuring Policies on Credit Risk Are All Mediated by the Risk Management Process, with Internal Control Acting as A Moderator, IJOQM, Volume 29, Number 3, September 2023, P62.
- 36 Guide on effective risk data aggregation and risk reporting, ECB, July 2023, P1.
- 37 Craig Davis et al, BCBS 239: Bracing for Change and Capturing the Opportunity, KPMG International Cooperative, 2016, P10.
- 38 Akber Dattoo, Legal Data for Banking: Business Optimisation and Regulatory Compliance, United Kingdom, Wiley, 2019, P86.
- 39 Irena Gecas-McCarthy et al, BIS assessment of BCBS 239 compliance, Deloitte, 2019, P3.
- 40 BCBS 268, Progress in adopting the principles for effective risk data aggregation and risk reporting, BIS, December 2013, P4.
- 41 Sean Smith et al, BCBS 239 Compliance: A catalyst for gaining competitive advantage, Deloitte, 2017, P5.
- 42 Richard Rosentha, Alexandre Brady, Cory Liepold, BCBS 239 Progress Report: Significant work still needs to be done for full compliance, The Deloitte Center for Regulatory Strategy, February 2024, P6
- 43 BCBSd599, Progress in adopting the Principles for effective risk data aggregation and risk reporting, BIS, November 2023, P 7.
- 44 Richard Rosentha, Alexandre Brady, Cory Liepold, BCBS 239 Progress Report: Significant work still needs to be done for full compliance, The Deloitte Center for Regulatory Strategy, February 2024, P7.
- 45 Talifhani Ramaliba, Lorette Jacobs, Artificial intelligence technology to enhance data quality management practices in the banking industry in South Africa, the South African Journal of Libraries and Information Science, 2024, 90(2), doi:10.7553/90-2-2392, P2.