2020

Data quality in the insurance sector

Stocktaking and proposed way forward

Table of contents

Executive summary

Data is a fundamental business and control asset for insurance undertakings, in particular to support informed decisions during challenging times. While the crisis triggered by COVID-19 is still unfolding, management actions are needed to cope with balance sheet volatility, better understand exposures, strengthen online channels, and optimize the product mix. Former insurance pricing schemes may be disappointing and insurance frauds are likely to spike, similar to what the industry experienced over past recessionary cycles[1]. More than ever, at the same time as the digital revolution keeps accelerating, good quality information is central for marketing, pricing, underwriting, reserving, management control, claims management and risk management.

Although insurance companies have always based their management decisions on information about exposures, risks and customers, the entry into force of Solvency II in 2016 has been a key incentive for insurers to progress from informal data quality management toward a more structured approach. Almost five years since the introduction of Solvency II, the CRO Forum has run a survey to take stock of data quality practices in the insurance sector.

In general, data quality refers to data's "fitness for use", which is the ability to fulfil the requirements of intended usage of data in a specific situation. Multiple data quality frameworks and approaches exist, although no one seems to cover all of the capabilities required to ensure an effective data quality implementation: objectives setting, planning, measurement, monitoring, organization, and tools management. As of today, data quality implementations in the insurance industry, typically:

- are owned by Risk Management and Finance;
- pertain to Solvency II internal models as well as financial reporting and reserving;
- characterize data according to attributes like appropriateness, completeness and accuracy.

1 McKinsey & Company, 2020

It is further observed that:

- data quality maturity outside the application of regulatory frameworks is not consistent;
- data quality is mostly handled at operational level, with little involvement of top management;
- lack of data quality is managed as an operational risk, but formal definitions of data quality risk tolerance and use of scenarios to estimate data quality risk are uncommon.

Data quality should start with objectives setting, planning, and process/system design. A major obstacle to objectives setting and, more in general, to a strategic approach to data quality, is the uncertainty around economic value of data. Data value can however be measured, for example using cost-based or income-based models. Investments in data quality, in particular, can be assessed through alternative cash flow scenarios.

Chief Risk Officers (CROs) can play a critical role in data quality management by:

- promoting a vision on data quality benefits;
- supporting the definition of optimal data quality governance;
- assessing data quality value and risk in the objectives setting phase;
- monitoring the actual data quality risk profile;
- helping business management to address data quality risk.

For a forward-looking approach to information and value-driven decision making, particularly in time of crisis, risk managers should join forces with business managers to develop analyses that evaluate investments in data quality.

1 Data quality, a growing need

Insurance companies experience data quality challenges on a daily basis. A wrong address leads to lost mail or a missing customer identification number turns the service centre call into a quest. Data quality plays a critical role in the success of an organisation, especially in a digital economy context, and can be a competitive advantage to those organisations that master it. Because of the digital connectivity of entire value networks, data errors and misuse are having more significant effects than they did in the age of isolated information technology applications.[2]

Financial assets, human resources, buildings, or machinery are fundamental assets to any organization. Data assets, however, have not gained similar attention from the management so far, even though the importance of data management has been emphasized since the 1980s. Data fulfils all of the characteristics of an "intangible asset" as defined by accounting standards, such as IAS 38 or IFRS. According to this definition, intangible assets are characterized as non-physical, separately identifiable, controllable, yielding an economic benefit when used, and capable of generating future benefit[3]. Moreover, data has unique characteristics that make it different from other assets. Data is easy to copy and transport, but it is not easy to reproduce if it is lost or destroyed. It is not consumed when used and can also be used for multiple purposes or by multiple people at the same time.[4]

Analysts can spend as much as 40% of their time validating data relevant to their analysis before any outcome can be used for strategic decisions[5]. At several insurance companies, executives are sceptical of data presented to them. Strategic thinkers will ask themselves about the cost-opportunity of lost customer insights or economic understanding. Little imagination is needed to see that radically different customer service propositions and operating models would be possible if data were of higher quality[6].

There are many different definitions for data quality, some very detailed, others quite technical. For the purpose of this paper, we define data quality "as a multidimensional construct that refers to data's *fitness for use*", namely the ability to fulfil risk management requirements in its processes[7].

Data quality is not only a necessary pre-requisite for effective risk management, but rather a risk in itself: an operational risk. Data quality is critical to successful processes, innovation[8] and the reliability of business reporting. Errors in data cause errors in reports generated from it. Lack of trust in data leads to wrong decisions and opportunities are missed when data is inaccurate, incomplete, delayed or incomprehensible.

2 Otto B., Österle H., 2015
3 Fadler M., Legner C., 2019
4 Sebastian-Coleman L., 2018
5 Forrester, 2019
6 Oliver Wyman, 2011
7 Juran J., 1974
8 Swiss Re, 2020

The high costs[9] of poor quality can be classified into process and opportunity costs: increased process costs, such as the costs associated with the re-execution of a process due to data errors, correction efforts and costs due to lost or missed revenues. That's why data quality should be guided and monitored in an effective data quality management system and ideally a "Quality at Source Philosophy" should be implemented.

But high quality of data is not a goal in itself. Companies using data of good quality have a better chance to reach their intended goals. Internal and external business requirements define the goals, and the quality of data is an essential factor to meeting such goals.

BUSINESS REQUIREMENTS DETERMINE THE GOALS

INTERNAL

- **Reliable corporate management**
- **Effective risk management**
- **Correct closing**

EXTERNAL

- **Compliance with legal and regulatory requirements.**
- **Information needs of rating agencies, business partners, analysts, shareholders, and other stakeholders**

What is actually poor data quality? "Wrong data" is usually the first swift answer. Almost every discussion on this topic starts with this simple consensus. But the real situation in business seems to be much more complex.

Data quality issues can be caused by certain areas in a company and manifest in other areas that use the data. Therefore, to make an efficient data management possible, it is essential to identify the problems in explicit terms. A data management framework would ensure that important data assets are formally managed throughout the enterprise and that data governance goals are achieved[10]. Good data management is the basis of an effective data quality management.

SPOTLIGHT

Data is an intangible asset, as:

- it is an identifiable and non-monetary asset without physical substance;
- it is likely to yield an economic benefit when used by the respective organization;
- the costs incurred for its production or procurement can be quantified;
- future economic benefits can be obtained and the access of others to these benefits can be restricted.

What do risk managers say about the importance of data quality:

Exposure
"The accuracy of any risk assessment is reliant on high-quality input data. Catastrophe models are particularly sensitive to poor data quality. Conversely, such models provide an opportunity to improve exposure data quality and, as such, enable risk managers to develop a comprehensive understanding of companies' exposures."

Reserving
"As empirical methods are the standard approach for P&C reserving, data quality is essential in this process. Data errors might end up in a wrong estimation of future financial obligations for those losses that are incurred but not reported. Furthermore, the calculated pattern - very important for the asset liability management - could be distorted. All estimations done in the reserving process could only be as good as the data quality."

Modelling
"All data from the treaty and exposure administration systems, which are relevant for the calculation of aggregates and thus for the utilization of risk tolerances, need to be checked for completeness and freedom from errors. This is crucial as the aggregate calculation delivers risk information data for the internal model, underwriting, supervisory authorities, rating agencies as well as for group retrocession."

Data types and data quality dimensions are terms that should be described as clearly as possible:

- Data types describe what kind of data should be analysed. A common language on data types makes discussions about data quality activities much easier.

- Data quality dimensions describe how these data types should be analysed. The regulatory perspective on the dimensions to be considered will be discussed in the next chapter, but several frameworks can be adopted[11].

Depending on the specific business or regulatory objective, different data quality dimensions can be critical success factors, but the functionality of the information system can only be ensured if all dimensions are of sufficient quality.

9 English L. P., 2009
10 DNV GL, Recommended Practice 0497
11 Rohweder, Kasten, Malzahn, Piro, Schmid, 2008

2 Requirements and regulations on data quality

Although insurance companies have always based their decisions on information about exposures, risks and customers, the entry into force of Solvency II Directive[12] in 2016 has been a key incentive for insurers to progress from informal data quality management toward a more structured approach. Solvency II is the first regulation that introduces strict requirements and detailed specifications for data quality for insurers.

Solvency II regime foresees data quality requirements in the following areas[13]:

Technical Provisions	Art. 19 to 21	Art. 34	Art. 264, 265	
USPs	Art. 219			
Internal Models	Art. 231 (1-3)	Art. 237	Art. 244, 245	Art. 247
Underwriting & Reserving	Art. 260			
Internal Control System	Art. 266	Art. 267 (4)		
Acturial Function	Art. 272			

The importance of data quality is also reflected in the situation where reported Solvency II data is used by National Authorities in the supervisory review process, by National Central Banks as input in the compilation of insurance statistics, as well as by EIOPA and the ECB for different market analyses[14].

12 The Solvency II Directive (2009/138/EC) is a Directive in European Union law that codifies and harmonises the EU insurance regulation. Primarily this concerns the amount of capital that EU insurance companies must hold to reduce the risk of insolvency. Solvency II came into effect on 1 January 2016.
13 Commission Delegated Regulation (EU) 2015/35 of 10 October 2014 supplementing Directive 2009/138/EC
14 Common Minimum Standards for Data Revisions as agreed between the ECB, EIOPA, National Central Banks and National Competent Authorities, 2019

Prudential regulation for the banking sector, Basel III[15] shows no major differences compared to Solvency II. The main data quality definitions are the same: accuracy, completeness and appropriateness. Furthermore, both regimes allow working with internal and external data and focus on sufficient documentation of used data, especially in case of data limitations. Only in respect to used length of data observation periods, Basel III is giving more and concrete details compared to Solvency II.

The General Data Protection Regulation (GDPR) is the regulation of the European Union that makes uniform rules for processing of personal data by most data processors, both private and public, across the EU. This is intended, on the one hand, to ensure protection of personal data within the European Union, and on the other hand to guarantee free movement of data within the European internal market. GDPR is the common data protection framework in the European Union since May 25, 2018. As insurance companies handle significant volumes of private data, GDPR plays an important role. For calculations or analysis, most of the data is already aggregated (anonymised) and therefore not underlying the GDPR. The quality of data is mentioned in “Article 47 (2.d): Binding corporate rules” only.

The International Financial Reporting Standards (IFRS) don’t foresee any specific data quality requirements but since it represents the core related to Accounting and Bookkeeping principles, it is clear that data quality is of high importance. Furthermore, financial results are used in many other processes of an insurance company (e.g. internal models, reporting, pricing).

Responsibility for the quality of Solvency II reporting (including the quality of reported data) usually rests with the CRO, the Chief Financial Officer or a comparable function. However, the ultimate responsibility for data quality should lie with the organisation as a whole, given the fact that data initiation and processing usually take place in various units throughout the organisation, such as insurance policy administration systems.

SPOTLIGHT

Data quality dimensions in the context of Solvency II

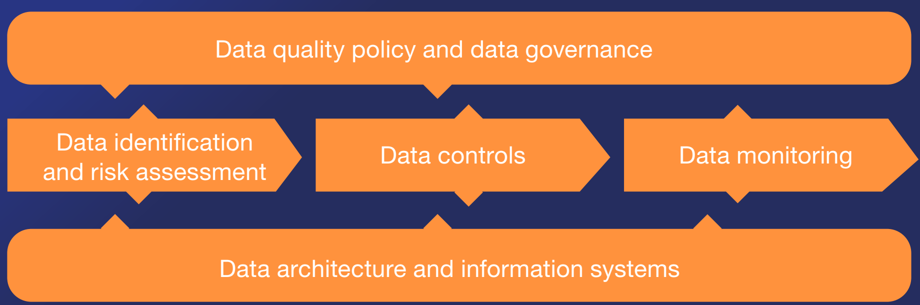
Accuracy
- Data is free from material errors.
- Data from different time periods used for the same estimation is consistent.
- Data is recorded in a timely manner and consistently over time.

Completeness
- Data includes sufficient historical information to assess the trend and characteristics of the underlying risk.
- Data is available for all of the relevant model parameters.
- Data is not excluded from use in an internal model without justification.

Appropriateness
- Data is consistent with the assumptions underlying the actuarial and statistical techniques applied in an internal model.
- Data reflects the relevant risks to which the undertaking is exposed.
- Data amount and nature ensure that estimations made in an internal model do not include material estimation errors.
- Data is collected and processed in a transparent and structured manner according to data definition and quality assessment criteria.
- Data is updated regularly or whenever circumstances command.

The elements of data quality management

The elements of the data quality management framework in the figure below may help give structure to insurance regulators’ expectations. A way to achieve this, is by defining various aspects of data quality management within the organisation, such as data governance, data quality, data management and data architecture[16].



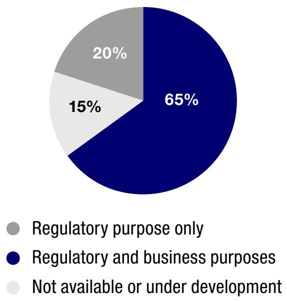
Data quality should:
- be a regular item on the Management Board’s agenda;
- receive ongoing priority attention within the organisation;
- be a structural component of operational management;
- be applied to the processes related to Solvency II reporting and also to the insurer’s operational processes in general.

15 Basel III is an internationally agreed set of measures, developed by the Basel Committee on Banking Supervision, aiming to strengthen the regulation, supervision and risk management of banks; Basel III was developed in response to the financial crisis of 2007-09.
16 De Nederlandsche Bank NV, 2017

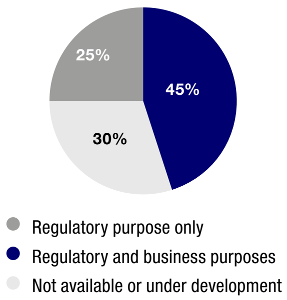
3 Findings on data quality status in the insurance sector

Almost five years since the entry into force of Solvency II, the CRO Forum has run a survey among its members to take stock on data quality practices in the insurance sector. Twenty members responded to 94 questions and provided a wealth of comments on data quality governance, scope of data quality application, practices and maturity. Key highlights are provided here.

Scope of data quality policy



Scope of data quality organisation



- Data quality may be governed in the context of a dedicated or broader policy.

- Regulatory compliance is a key motivation to formally manage data quality.

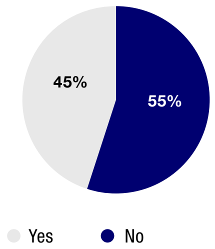
- Prevailing regulatory focus is on Solvency II, financial reporting and reserving, but data protection (GDPR) regulation is also relevant.

- Business processes like customer relationship management and management control may be managed outside the scope of formal data quality governance.

- Only 50% of respondents have a formal data governance committee, but in many cases other committees take care also of data.

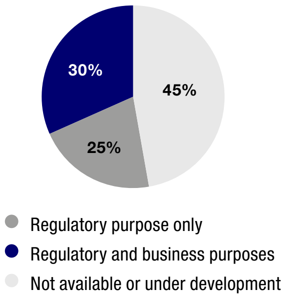
- 65% of respondents are planning to further develop the data quality organization.

Periodic review of data management objectives



• The somewhat static approach to objectives setting by the majority of respondents, suggests that data quality governance may not always be aligned with the undertaking's strategic objectives.

Availability of data assets maps



• When available, maps of data assets are often embedded in data directories.
• Maps are typically available for a certain scope.
• GDPR has been a key motivation to develop data assets maps.
• Documentation on single sources of truth for specific data is generally limited.

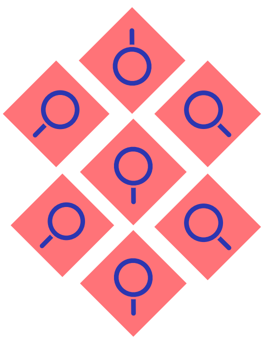
SPOTLIGHT

Use case – Data quality for Solvency II

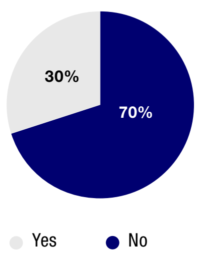
A CRO Forum member has implemented a group-wide data governance framework. A data quality policy and data quality standards are in place, while a group Chief Data Quality Officer coordinates divisional data teams, and Finance and Risk Management departments play a key role. Catalogues for data assets, flows, risks and controls are hosted by dedicated tools and updated on a regular basis. Data assets are categorized by criticality according to expert judgement and quantitative tests. Since 2016, the scope of data governance includes reserving and Solvency II internal model. A data quality risk tolerance is defined based on the materiality principle with regard to Solvency Capital Requirement and technical provisions. Data quality risk is evaluated by actuaries and risk experts at legal entity level by means of sensitivity analyses. Risk drivers include data incompleteness, incorrectness and inappropriateness as well as control deficiencies. Data quality risk reporting is shared regularly with top management. Any deficiencies are handled through a follow-up and escalation process. The internal model validation team ensures that model changes are processed into the data quality management system.

Use case – Data quality for personal data

A CRO Forum member has Data Governance and Data Privacy group policies in place. The Data Privacy policy requires that personal data about customers, intermediaries and personnel are accurate, consistent and updated. The group framework is cascaded to business units. A maturity model with a one-to-five scale is adopted to define local implementation minimum standards. Critical data sources and systems have been identified based on Data Owners' expert judgment and also based upon the categorization of assets made in the context of business continuity management. Data risks and controls catalogues are refreshed on a regular basis and hosted by the operational risk management tool (data risk management is part of operational risk management). Group-wide data governance and data privacy risks have a defined risk tolerance of their own. For data privacy and the management of personal data, the risk tolerance is further defined according to a quali-quantitative approach. The group sets an overall tolerance level that business units may make stricter. Acceptability thresholds have been defined with respect to financial losses, reputational impact and customer detriment. These thresholds are monitored through the execution of about 25 data quality validation rules, each calculating a percentage of valid data records according to a specific criterion. Tolerance is monitored twice-a-year and results are summarized into a dashboard. Issues are first reported to business units' Data Owners and, if necessary, escalated to business units' data forums and then risk committees.



CDO role established

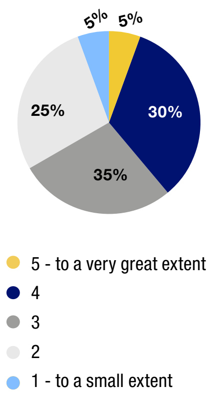


• Other roles than the Chief Data Officer (CDO) often take on the responsibility for structuring and managing data.

• When established, the CDO typically reports to the head of another domain like IT or Risk Management, and can either manage or oversee the data quality risk (50%/50%).

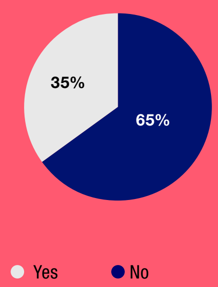
• The CDO can be either a doer or facilitator.

Level of CRO involvement



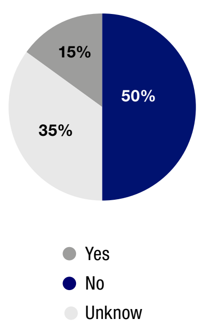
- CROs enacts the second line of defence and sponsor high quality data.
- The CRO can also define the data governance policy, chair the data quality committee, identify critical data fields and monitor the quality of data.
- Data quality is typically handled within the operational risk management framework.
- The level of CRO involvement varies across organizations.
- Specific reference to the management of data value is infrequent.

Data quality risk evaluated in economic terms



- Although most respondents perform formal data quality risk assessments, only 35% evaluate risk in economic terms and have defined risk tolerance quantitatively.
- Data quality risk tolerance can be defined based on experience or calibrated based on risk impact and likelihood.
- Specific stresses and scenarios are used only by a handful of respondents to evaluate data quality risk.
- Data quality indicators are used by about half of the respondents.

Data quality metrics reported to CxO or board



- Although most respondents implement formal communication and reporting (covering both risks and losses), data quality metrics are only rarely shared with senior management.
- Data quality is mostly handled at operational level.

SPOTLIGHT

Data quality challenges in the banking sector

In May-June 2018, the European Central Bank (ECB) and the Basel Committee on Banking Supervision (BCBS) published reports on the progress of the largest, internationally active banks towards compliance with the BCBS Principles for Effective Risk Data Aggregation and Reporting - known as BCBS 239.[17] Whilst reports approach the topic from different angles, similarities in their findings are striking and paint a scenario where, two years after the original compliance deadline, gaps are still significant and widespread. Observations made by both ECB and BCBS focus mainly on two areas:

1. Governance

- Banks show weak governance arrangements around data aggregation and reporting capabilities.
- Roles and responsibilities are often poorly defined and enforced between different functions.
- There is a lack of strategic attention to data at executive and senior management level.

2. IT Infrastructure

- Banks are still over-reliant on manual processes and siloed IT systems for risk reporting, hampering their ability to aggregate data.
- Cumbersome and complex reconciliation processes are used as compensating controls for poor data flows and infrastructure.
- Banks are often unable to generate reports in a timely manner due to underlying infrastructure and data flow issues - especially under stress scenarios.

Findings of the GIRO (General Insurance Research Organization) Data Quality Working Party

In order to examine the effect of data quality problems on critical financial quantities, the Working Party conducted a data quality experiment with actual data used for an actuarial application.[18] This experiment was designed to examine the effect of incomplete and/or erroneous data on loss reserve estimates. Real loss triangle data was felt to be more persuasive than conducting the experiment on a simulated dataset. Data of sufficient maturity were obtained - all years were fully developed and the true ultimate losses were known - and various methods were employed to estimate ultimate losses using the data as of past valuation dates.

One of the data challenges that practicing actuaries frequently encounter, relates to datasets that are severely limited with respect to the completeness of information provided. That is, the data may be limited with respect to the numbers of years of history (e.g. only five years of history for a long tail line where claims take 20 years to fully settle) or the types of data provided (e.g. only paid and incurred losses, but no reported claim count, closed claim count or exposure data). Another data quality challenge that the working party investigated was data accuracy. Modifications were intentionally introduced into the data to simulate data errors and data quality problems commonly encountered.

The outcome of the data experiment indicated that there was a significant increase in the uncertainty of results and a significant decrease in the accuracy of results when data quality problems were present. The errors resulting from poor data can significantly reduce the reliability of actuarial analyses, and this could have a direct effect on an insurer's financial statements. The GIRO Working Party stated that insurers should devote more time and resources to increasing accuracy and completeness of their data by improving their practices for collecting and handling data. In particular, insurers would benefit from the investment of increased senior management time in this area. By taking such actions, they could improve their efficiency and hence their profitability.

17 Deloitte, 2018
18 Casualty Actuarial Society, 2008

4 Data quality management best practices

Data quality policy

Having a data quality policy in place with group-wide principles/objectives is the first step to the successful implementation of data quality management: specifying the reason why, the scope and the approach. The extent to which the data quality policy is defined on a central or decentral level, depends on the specific organisational structure. In a changing environment it is important to periodically review the data quality policy at least once a year.

Data assets

A map of data assets is often documented in a data dictionary, where data is specified including the required metadata[19]. In general, three types of metadata are defined: business metadata (e.g. definition, data classification, data quality rules), technical metadata (e.g. technical name, data type, location) and governance metadata (e.g. Data Owner, Data Steward). Some regulations, like Solvency II, prescribe the scope of data (key data) based on the impact of the individual data on regulatory reports (e.g. provisions, SCR and MCR) using materiality and risk appetite. Building the data lineage (following the data from source to report) is essential to identify all data involved in the creation of a report. The accurate, complete and appropriate journey of data from source to reports (lineage) should be guaranteed by process level controls like reconciliations.

Data quality organization

Data quality organization consists of roles at an enterprise level (for setting standards and helping the local entities in data quality matters) and roles within local entities. To keep alignment between these roles, it is necessary to have communication structures in place, for example committees. In general, there are committees at three levels:

- Strategic level, also known as Data Governance Board, including the data sponsor/CDO and representatives from departments involved in data management (mainly CRO, Actuarial, Finance, Internal Controlling, Information Technology, Chief Information Security Officer and Data Privacy Officer – the Board of Directors has to be regularly involved in data quality investments).
- Tactical level, including all Data Owners for one entity.
- Operational level, including all Data Stewards for one entity.

19 Metadata is “data about data” and describes the properties of other data. Typically, metadata enables retrieval and maintenance of data containers (e.g. documents or files).

A dedicated team or function is required to structure and maintain data quality organisation and committees. CDOs may have a more strategic or security mission; they can also facilitate the implementation of data quality policy and supervise adherence to procedures in cooperation with e.g. the Chief Information Security Officer, Data Privacy Officer and CRO. The CDO may be the main contact for internal and external stakeholders regarding data quality matters/issues, sign-off on data quality reports and chair the Data Governance Board.

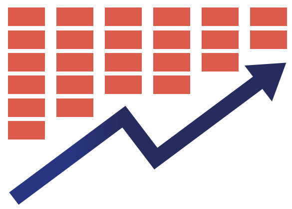
Good data governance requires clear objectives and roles (e.g. Data Owners, Data Stewards, Data Users and Data Custodians can be defined). Data ownership is applied to the creation or origination of data where Data Owners are accountable for the quality of their data and for empowering their teams to perform work. The Data Steward, in many cases, is most knowledgeable about the data and therefore plays an important role in setting the definition and metadata, setting data quality rules and monitoring data quality.

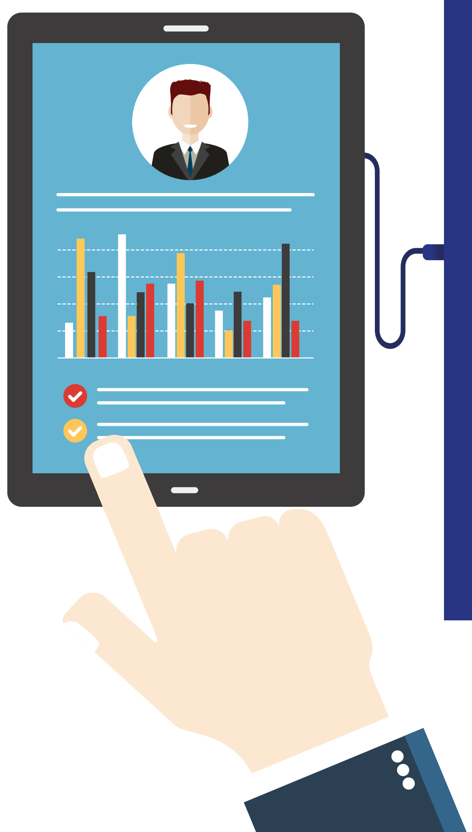
The role of the CRO or Risk Management function is key in policy setting, defining the data scope (key data) and monitoring data quality. The Compliance function checks the adherence to legislation and prepares policies, it complements the second line of defence for data quality monitoring and acts upon data quality compliance issues. There is a close cooperation between IT department and data management team, where there is a strong interdependence. IT department is involved in the implementation of business requirements and supports the implementation of technical and automated controls, and data related tooling (e.g. data governance tooling and data quality tooling). For applications that remain within the remit of functional or business divisions, IT security and operability standards are still applied.

To enable data professionals like Data Owners and Data Stewards to perform their work properly, trainings must be available. Data awareness trainings can bring the organisation to a certain level of data literacy. Besides that, employees with a specific role should be trained in their tasks and responsibilities. To ensure the fulfilment of responsibilities, it is desirable to make data quality part of the review cycle.

Tooling

Tooling is supportive of data quality management activities. Functionalities expected in a data quality tool are directories/repositories (for data/ metadata/controls), controls execution and outcome collection, data quality reporting, exception handling, data/ process mapping, data extraction/ transformation, data profiling, data lineage and audit trails. Artificial intelligence can be integrated to detect and help resolve data quality issues that traditional systems would not identify[20].





SPOTLIGHT

Data Quality Controlling System: summary and best practices

Data quality is typically initiated and based upon regulatory requirements and data quality controlling frameworks or data governance frameworks (e.g. Data Governance Institute's Framework, Data Management Association's Body of Knowledge). Among regulatory requirements, Solvency II and Sarbanes Oxley control frameworks are common. Six Sigma process quality systems are also relevant.

Data should be controlled throughout production, storage and processing. Automated and manual controls ought to be positioned on data flows, with acceptance thresholds. In case of manual checks, those responsible for carrying out these checks, should be automatically reminded. Regular clearance reviews can guarantee proper authorizations management.

A control plan should be formalized and kept up to date for a regular review of all key controls. Data quality risk identification can be done based on a materiality assessment or based on expert judgement. Data quality risk can be assessed separately or as part of an Operational Risk Management (ORM) framework or Integrated Risk and Control System (IRCS). A data quality risk taxonomy can be based, for example, on Solvency II or Financial Reporting Risk frameworks. Incident registration and actual loss assessments are generally registered in the regular incident management systems.

Probabilities and impacts are often estimated based on expert judgement. Experts have the possibility to use scenarios with different types of consequences to support their estimations. Financial impact for data quality is generally evaluated as part of the Value at Risk calculation. Data quality risk tolerances are determined in multiple ways, for example based on specific data quality criteria, reporting thresholds, materiality frameworks, expert judgement and potential impact/likelihood set. Data quality indicators are used by management to monitor data quality. Examples of data quality indicators are a measure of the percentage of completeness or the number of failed data attributes.

20 Fadler, Legner, 2019

SPOTLIGHT

Data quality rules

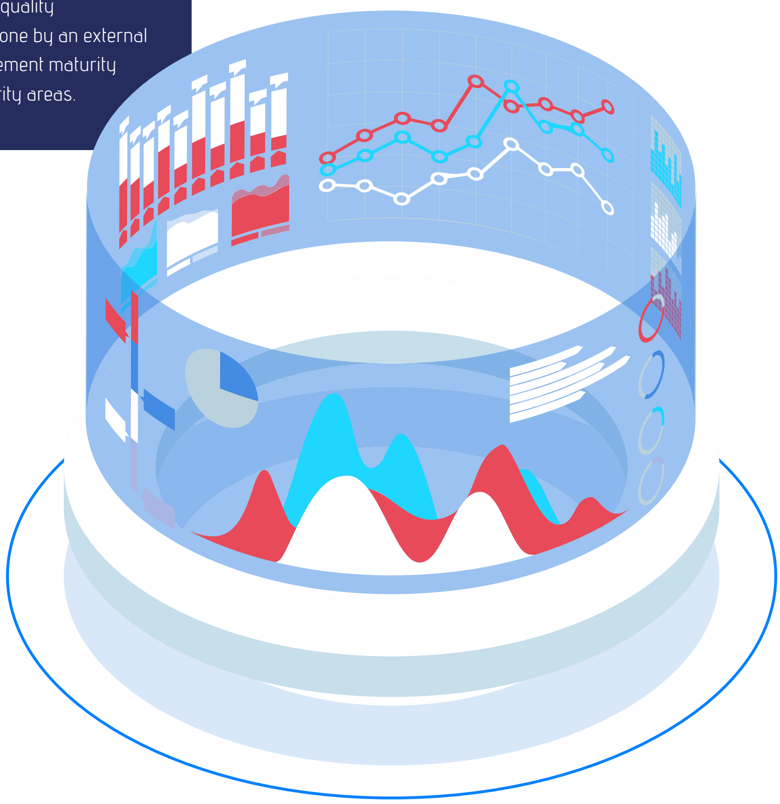
Data quality rules[21], also known as data validation rules, define the business requirements for specific data and can be used to check the quality of data and data records. They are the primary tools for determining data quality. By checking against validation rules, it is possible to test whether data meets the defined criteria and possesses the required attributes. In this way, potential weak points (e.g. in processes) can be detected and recommendations for action can be derived. Data quality rules allow for the measurement of different data quality dimensions, such as completeness and accuracy. [22]

Attribute	Sample Rule
Completeness	The Line of Business entry may not be empty.
Accuracy	The Date entry must have the proper format.
	The value in the Premium entry may not be negative

A ready-to-use data quality rule is an algorithm that checks the structure, format and arrangement of data. However, data validation rules are generally conceptualized and described prior to developing an executable, machine-readable validation rule.

Maturity assessment: summary and best practices

It is advised to periodically perform an assessment of the data quality management environment. This can be a self-assessment or done by an external party. Frameworks that can be used are common data management maturity models or own models. Results should be used to identify priority areas.



21 Schlosser S., 2015
22 Otto B., Österle H., 2015

5 Enhancing the data quality opportunity

Given the centrality of information to insurers, it may seem surprising that not all organizations explicitly recognize the issues associated with poor data quality. Some organizations have learned to work around the issues, often resulting in sub-optimal reporting and/or decision making. Executives find fixing data quality difficult, expensive and time consuming.

Data quality programmes typically start with a framework consisting of measurement, governance, ownership/stewardship, processes and organisation. Such a framework must consider, as a minimum, a consistent data dictionary, data lineage, a data ownership model, data quality criteria, and clear governance over data changes. While these are undeniably important, implementing such a framework often fails to truly improve information because it is too hard to implement, especially if key internal stakeholders do not commit.

If you can't measure it, you can't manage it

Despite the increasing importance of data, data assets rarely receive the same level of attention as other organizational assets. One reason for this reluctance may be the lack of broadly accepted methods for data assets valuation. Still, recent studies have shown that monetary value of a data asset can be assessed using theoretical approaches for valuation of intangible assets. This value, called "value-in-use", depends on the way data assets are used in business processes within a specific organization.[23]
Although the value-in-use cannot be interpreted as an objective financial value arising from market prices or the cost of producing the data, the method has noteworthy outcomes:

- A financial value of the performance of data assets.
- Quantified opportunities and risks as a result of using data in certain data use contexts depending on the given level of data quality.
- Decision-making criteria regarding increased process efficiency and data quality improvement.
- Transparency regarding data quality management and performance of data assets.
- Intra-organizational benchmarking regarding data use efficiency.

23 Zechmann A., 2016

In determining cost savings of improving data quality, one should ideally review the information flow through the business process to determine where, in the process, the error is introduced. With research, one might be able to consider the cumulative costs related to business impacts over the number of times an error occurs.

Once the source of error introduction is identified, the data analyst can consider alternatives for eliminating the root cause, instituting preventative techniques, and/or taking some corrective action. Each, or perhaps all, of these alternatives require an investment of both money and resources for acquiring any appropriate technologies, staffing for designing, developing, and implementing solutions, training, and ongoing maintenance of the solution. Providing an estimate at this point establishes a cost for remediation.[24]

	MARKET APPROACH	COST APPROACH	INCOME APPROACH
BASIC IDEA	The value of an object can be derived from the value of identical or similar objects traded in the market	The value of an object equals the cost incurred for making or buying an exact copy of it	The value of an object equals the total economic benefit created by the object in the future (net present value)
MAIN VALUE DETRMINANTS	• Market prices • Analogous prices • Market-to-book ratios • Multipliers	• Reproduction cost • Replacement cost	• Revenue increase • Cost savings • Capital requirement reduction
APPLICABILITY FOR DATA VALUATION	Often not applicable, as the precondition of active data markets and transaction prices is often not given	Efficiently applicable, but has its limitations, as it is not future-oriented and does not take the benefit created by data into account	Applicable, but requires substantial effort (e.g. making forecasts about future cashflows created by data)

Approaches for data valuation (based on Zechmann, 2016)

SPOTLIGHT

Striking the right balance

“When you take a lock down approach to information flow and you over-govern it, you destroy value, and if you under-govern it, you won’t maximize the value to the business. It’s an interesting balancing act and one that we constantly have to monitor and adjust. If we just focus on protection, we will over-control and constrain and then we’ll generate actually more risk since people will go around the controls. The only way to manage risk is to run towards it and be at the front of it. That’s the point of opportunity rather than just trying to contain it. So, we protect to enable. If we don’t, then we’re not in the sweet spot of reasonable risk and reasonable controls to maximize the value of information assets to the firm. If you constrain, you drive the business to create. You will drive users to go out and do it on their own, which without the right expertise, is like a kid running with scissors. Unless you can somehow manage that, you will have ungoverned data storage that will create more challenges down the road, more risk, more compliance issues, and probably a far higher total cost of ownership.”

(Chief Information Security Officer, Intel Corporation)[25]

24 Loshin D., 2011
25 Tallon P. Ramirez R., 2013

The human perspective should also be considered
The organization must empower employees to perform roles and fulfil responsibilities to achieve governance goals related to data quality, which may require re-training of staff. Effective governance includes having key performance indicators on data quality improvements, better transparency on data status and, most importantly, backing by senior management.[26]

Building data quality into the DNA of an organisation is not a short-term project that can be undertaken within any given operating function. Instead, radical changes to the culture, priorities and incentives of top-level executives down to call centre staff are required. Buy-in from the Executive Committee is necessary to ensure that the "life or death" influence that data quality has on strategy and operations is recognised. What can executives actually "do" differently? The following are critical:

• Build a case for action by quantifying the cost and/or opportunity of data. FTE costs of manual data reviews should be visible.

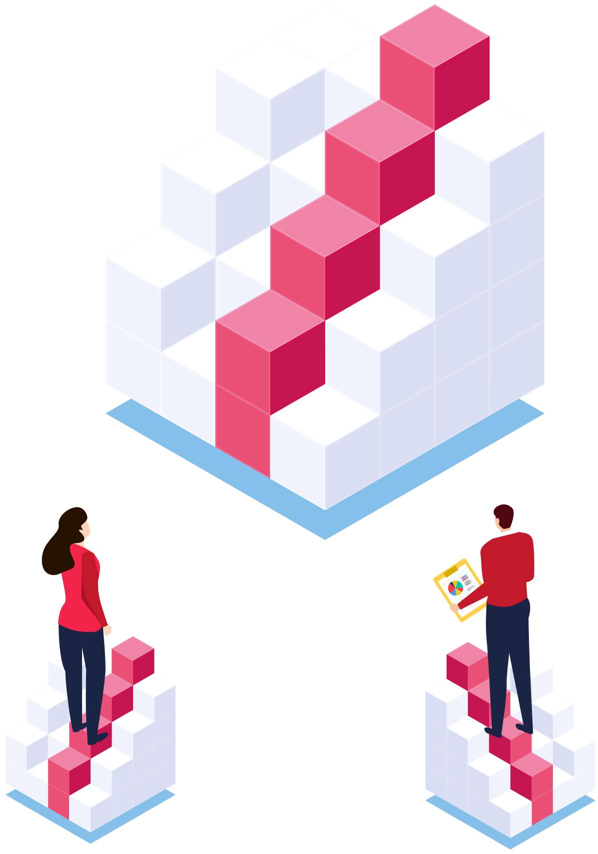
• Embrace the challenge: critical issues pervasive to organisations are governed day-to-day by executives and data quality should be no exception. If group-wide financial planning is worthy of ongoing executive attention, then so is data quality.

• Tackle incentives head on: data quality is an incentives problem – for individuals that impact data quality (i.e. most executives and a majority of staff members), data quality should appear in incentive schemes, job descriptions and performance evaluation.

• Articulate what good looks like: most organisations have several data quality definitions occupying dusty drawers. Internal marketing of a business-relevant definition helps 'sell' the issue and lays the groundwork for other initiatives.

• Maintaining focus on a small set of critical business decisions and proving that success is possible help breaking the failure cycle.

• Instil discipline over "proprietary" data: in a data quality vacuum, many functions in an insurance company will have their own "proprietary" version of the truth, undermining transparency and access to information. Phasing out "proprietary" data has to be done carefully (it was created for a reason), but executives should set and stick to hard standards for data sources to meet, and these should include a single version of the truth. [27]



SPOTLIGHT

Organizations often suffer from complex IT infrastructures, siloed applications and multiple disconnected processes that weaken the deployment of data quality initiatives. Overcoming IT infrastructure issues, the disparate quality of systems and data sources, and integrating legacy and big data environments, is one of the main challenges to ensure a successful deployment of data quality initiatives.

Master Data Management (MDM) is a well-known data management discipline. The consensus-driven definitions of business entities and their consistent application across an enterprise are critical success factors for important cross functional business activities, such as business intelligence, complete views of customers, operational excellence, regulatory reporting, compliance, mergers and acquisitions, and treating data as an enterprise asset. Most MDM programs focus initially on the customer data domain (MDM for a single application can be a safe and effective start), and then they move on to other domains, like products, financials, partners, employees, and locations.[28]

26 DNV GL, Recommended Practice 0497
27 Oliver Wyman, 2011
28 Russom P., 2012

6 Contribution of Risk Management

Improvements in data quality are driven by a variety of factors. Regulation and smoother operations are major drivers for insurance companies to adopt risk management in data governance.

The other leading driver is the need to generate new business and increase profits. These multiple drivers need a blend of data management, risk management and data governance principles, which raises several questions:

- Is risk management an integral component of your data governance? Does your risk governance consider data management principles? How do these work together?

- Does data risk management align with Data Governance division or Chief Risk Office?

- Have you considered addressing governance gaps and overlaps between risk and data functions?

- Are data quality risks considered as a priority to your organization and have you cascaded these risks to your data governance operational frameworks to reflect these priorities?

- How do you classify data and how is it managed depending on its value?

Common enterprise data quality view
Risk management needs to be an integral dimension of data governance, for which policy needs are to be defined in close coordination with the enterprise risk function. The CRO should partner with the relevant stakeholders (e.g. Chief Information Officer, CDO and Data Privacy Officer) to revise policies, establish data quality standards and develop KRIs for measuring and monitoring data risks. The CRO should include measures of data quality in existing control frameworks to assure continuity of data governance measures and that key actions are developed and regularly monitored and reviewed.

It is necessary, for the CRO, that the different stakeholders of the enterprise share the same view on data quality risks and objectives in order to ensure that current data initiatives are aligned with the company business objectives.

This common vision could be achieved through a data governance framework that must encompass the organizational structures needed to achieve the required level of data quality. This includes the Data Governance Committee or similar, roles as Data Owners, Data Stewards, Data Custodians or similar.

In a highly regulated industry, Risk Management teams should be cross-functional, working with colleagues in multiple business lines and departments to assess regulatory requirements and identify the most cost-effective and efficient methods to deliver necessary data in an accurate and timely fashion. This will also ease the sharing of the same vision on data quality risks.

Senior management support, sponsorship and understanding
Making senior leadership support, sponsor, and understand the activities of data governance and results of governing data, is a must. It is a challenge to get senior leadership to do all three. Effective communication is necessary. This could be achieved with the help of the CRO by highlighting the risks of low data quality and providing estimates on the costs of bad data quality and non-compliance to regulations.

Roles and responsibilities
Organizations often struggle to define the best place for data quality roles. The CRO should make sure that roles of the different units, internal bodies and staff involved in the data quality management process, are defined in such a way as to ensure that the data handling process is sufficiently independent from the data quality management process to guarantee an effective segregation of duties. One of the best practices is to have a dedicated independent unit with an overall view and responsibility for the management of data quality. The accountability for reviewing the alignment and adherence of data governance risks with enterprise objectives, can be assigned to a standalone function within the Risk Office.

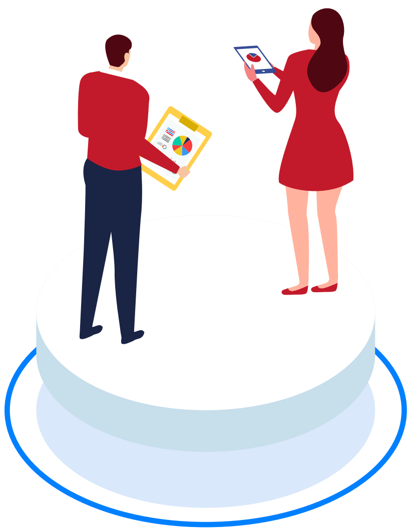
Risk-based approach
Data quality management should be based on management and governance needs, and these vary as the scope changes. All defined processes, policies and procedures should comply with and adhere to the overarching enterprise data governance program. The CRO should help identify key processes that possess critical data as prioritized areas to implement data quality. Areas that can be classified as "high risk" based on:

- highly regulated processes;
- high volume of data transactions;
- technically complex performance information;
- previously identified problems;
- inexperienced staff involved in data processing;
- a system being used to produce new performance information;
- known gaps in the control environment.

The degree to which an organization can use data strategically, is the degree to which data is effectively governed. The next challenge, once scope is defined, is the definition of the threshold for good enough data quality as well as clear and quantifiable goals. The use of business KPIs and risk appetite figures, which can be provided by the risk office, will support the definition of specific quantitative guidelines in relation to acceptable data quality levels.

Budgets, ownership and staffing
Budget is one of the challenges that most organizations face. To overcome this, CROs should align their data quality efforts with key business initiatives and high risks mitigation. Defining metrics and thresholds, and estimating potential benefits, are key. To deliver on data quality objectives, CROs also need to ensure that the right resources and skills are available or planned for.

Costs and benefits
Quantifying the cost and/or opportunity of data quality is not an easy task. While quantifying direct costs of low data quality, like FTE costs of manual data reviews, or some non-compliance costs will be generally manageable, estimating the indirect costs of low data quality (e.g. lower reputation, wrong decisions or actions, sunk investments, productivity decrease) is way more complicated and hazardous. CROs should, therefore, be able to identify and retrieve the amount of losses due to poor data quality and further build risk scenarios that quantify bad quality risks in order to help establish a view on the cost of non-quality. This could be leveraged to build business cases or help prioritize areas depending on the risk appetite. Given the increasing regulatory requirements around data quality and the possible high sanctions (up to 4% of the whole company turnover for a breach of the GDPR), business cases should account for the cost of non-compliance.



SPOTLIGHT

CRO's role towards data quality

- Promote a vision on data quality benefits for an organization.
- Act as a critical enabler on strategy setting and project portfolio definition of data quality initiatives.
- Develop data risk scenarios including data quality risks for stress testing and capital planning activities.
- Identify legal, regulatory and contractual requirements, and organizational policies and standards related to data quality, to determine their potential impact on business objectives.
- Use business KPIs and risk appetite to define data quality standards and thresholds to identify and prioritize mitigating actions regarding data quality risks; then establish the data quality policy and standard to be used to drive behavioural change in business areas.

Value proposition for executives to gain support

Although certain costs like those of angry customers and bad decisions resulting from bad data are hard to estimate, a risk-based approach to data management provides the executive suite with clear and compelling benefits. For Chief Executive Officers focused on building trust with stakeholders/customers, adhering to regulations, increasing revenues, and lowering operational costs, a risk-based approach can improve decision making and trusted reporting.

For Chief Financial Officers concerned with ensuring accurate financial and management reporting, a risk-based data quality program establishes the policies, procedures, controls and measurements necessary to improve accountability, reporting and performance. For Chief Information Officers seeking to optimize information technology investments, a risk-based data management program improves applications' effectiveness, lowers infrastructure costs and, in some cases, establishes the ROI-based business justification that accelerates consensus and projects approval.

Conclusions

With this paper, the CRO Forum takes stock of data quality practices in the insurance sector. Formal data quality frameworks are implemented primarily to comply with regulations, while data quality management for business uses is still somewhat based on customized approaches. High quality information, however, is a prerequisite for obtaining customers' trust, running smooth operations and making business profitable.

Despite the increasing importance of data, also consequent to the digital revolution, data assets rarely receive the same level of attention as other organizational assets. One reason for this is the lack of well-established methods for data assets valuation, albeit recent studies have shown that the value of data can be measured using approaches for the evaluation of intangible assets. These methods quantify opportunities and risks stemming from the use of data in a certain context, depending on the given level of data quality. Thus, any undertaking's first step in the data management journey should be to articulate a data strategy and align data initiatives with business objectives based on value creation.

Since business and regulation evolve more rapidly than systems and processes, data assets tend not to be organized to properly support use cases. Moreover, data quality is rarely considered from the project design phase and controls are normally detective, rather than preventative. From the non-embedding of data quality in the design phase, it can follow that fixing data quality can be difficult and expensive. A cultural change towards data quality-by-design is beneficial.

Data quality programmes typically start with a framework consisting of governance, ownership/stewardship, processes and organisation. Although these are undeniably important, implementing such a framework often fails to truly improve information because it is too hard to apply. Striking the right balance between under-governing data and pushing the organization to work around uncompromising rules, can be difficult. The right solution depends on the corporate-specific organization and culture.

In line with the traditional definition of the risk management role, CROs' desirable contribution to successful data management includes interpreting regulatory requirements, helping the establishment of an optimal governance, supporting the definition of a risk appetite, identifying critical data, and organizing an effective data quality controlling system. Nevertheless, risk management's essence is to protect and enhance value generation. To truly integrate risk management into data management, CROs should partner with senior business executives to establish value driven data management, by evaluating data use scenarios.

Bibliography

Accenture, Improving Business Decisions While Addressing a Key Solvency II Requirement, 2012

Casualty Actuarial Society E-Forum, Dirty Data on Both Sides of the Pond, 2008

Commission Delegated Regulation (EU) 2015/35 of 10 October 2014 supplementing Directive 2009/138/EC

Common Minimum Standards for Data Revisions as agreed between the ECB, EIOPA, National Central Banks and National Competent Authorities, 2019

De Nederlandsche Bank NV, Solvency II data quality management by insurers, 2017

Deloitte, BCBS 239 – The "Data-Fitness" Hurdle, 2018

DNV GL, Recommended Practice 0497

English L. P., MIT Information Quality Industry Symposium, July 15-17, 2009

Fadler M., Legner C., Managing data as an asset with the help of Artificial Intelligence, 2019

Forrester: The Data Management Playbook for 2020 (2019)

Juran J., The Quality Control Handbook, McGraw-Hill, New York, 1974

Loshin D., Improved Risk Management via Data Quality Improvement, 2011

McKinsey & Company, Insurance in a rapidly changing coronavirus world, 2020

Oliver Wyman, Data Quality: The Truth Isn't Out There, 2011

Oliver Wyman, Imagine If Your Data Were Good, 2015

Otto B., Österle H., Corporate Data Quality Prerequisite for Successful Business Models, 2015

Rohweder, Kasten, Malzahn, Piro, Schmid "Informationsqualität – Definitionen, Dimensionen und Begriffe" in Hildebrand et al. „Daten und Informationsqualität", Wiesbaden 2008, P. 25 ff.

Russom P., Next Generation Master Data Management, TDWI Best Practices Report, 2012

Schlosser S., Data quality rules, 2015

Sebastian-Coleman L., Navigating the Labyrinth, TechnicsPub, 2018

Swiss Re Institute, Data-driven insurance: ready for the next frontier?, 2020

Tallon P., Ramirez R., The Information Artifact in IT Governance: Toward a Theory of Information Governance, Journal of Management Information Systems, December 2013

Zechmann A., Assessing the Economic Value of Data Assets, Work Report 2016