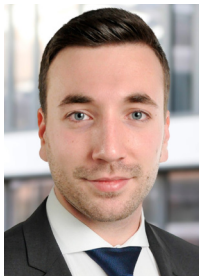
# Datenqualität in Banken – der Schlüssel für mehr Compliance

*„Data is the new Oil".* Diese Metapher ist inzwischen sehr verbreitet, weil sie sehr zutreffend den Bedeutungszuwachs beschreibt, den Daten im Rahmen der fortschreitenden Digitalisierung und Vernetzung gewonnen haben. Strenge regulatorische Vorschriften hinsichtlich Compliance rücken die Datenqualität immer stärker in das Blickfeld der Aufsicht.

*Autoren:*
*Christian Behrens,* Managing Consultant bei der movisco AG in Frankfurt

*Michael Junklewitz,* Senior Consultant bei der movisco AG in Hamburg



*Robin Müller,* Consultant bei der movisco AG in Frankfurt

*Die Auswirkungen von Compliance-Forderungen haben direkten Einfluss aufs Geschäft*

Operativ ist die Datenqualität ebenfalls ein Erfolgsfaktor, denn nur eine hohe Qualität der gespeicherten Informationen bietet die Grundlage dafür, dass Banken Risiken rechtzeitig erkennen und aktive Risikosteuerung betreiben können.

## Herausforderungen Datenmanagement

Der Finanzsektor sieht sich aktuell großen wirtschaftlichen Herausforderungen ausgesetzt. Geändertes Kundenverhalten, hoher Kostendruck bei einem gleichzeitig schwierigen Markt- und Zinsumfeld sind Rahmenbedingungen, auf die es strategisch und operativ Antworten zu finden gilt. Gleichzeitig hat das Management strenge regulatorische Rahmenbedingungen zu erfüllen. Zur Einhaltung der Compliance des Geschäftsbetriebs seien an dieser Stelle die strengen Vorgaben innerhalb der IT hervorgehoben.

Dazu gehören etwa:

- die Bankenaufsichtlichen Anforderungen an die IT (BAIT),
- die von der „European Banking Authority" herausgegebene Leitlinie „Guidelines on internal governance" (EBA/GL/2017),

- die Mindestanforderungen an das Risikomanagement (MaRisk) und
- das vom Basel Committee on Banking Supervision (BCBS) herausgegebene Papier #239 (Grundsätze für die effektive Aggregation von Risikodaten und die Risikoberichterstattung)

Die Compliance-Forderungen haben unmittelbaren Einfluss auf das Tagesgeschäft innerhalb des Instituts. Mit den oben genannten Vorschriften sind die Anforderungen an die Qualität, Umfang und Auswertbarkeit des Datenhaushalts gestiegen, was sich direkt auf das Reporting der Banken auswirkt. Aktuell deutet nichts darauf hin, dass dieser Trend sich wieder umkehren könnte. Die Digitalisierung und damit die Veränderung von Geschäftsmodellen und Prozessen führen parallel dazu, dass immer mehr Datenhaushalte mit exponentiell steigender Datenmenge entstehen. Das wiederum erschwert die Aufgabe für das Berichtswesen, die geforderten Informationen in der geforderten Zeit aufzubereiten.

Die wachsende Quantität an Daten innerhalb einer Bank führt ganz automatisch zur Frage, wie es um die Qualität dieser Daten bestellt ist. Die Notwendigkeit möglichst korrekte Daten zu verwenden, erwächst konkret aus zwei Erfordernissen:

1. Nur so werden die regulatorischen Anforderungen der Aufsichtsbehörden befriedigt.
2. Korrekte Daten und Informationen bilden die Grundlage für eine optimale Risikosteuerung.

Die Einführung eines Datenqualitätsmanagements (DQM) ist ein wesentlicher Schritt auf dem Weg, konsistente Daten für Berichte zu erhalten, und damit einerseits die Compliance einzuhalten, andererseits auch eine größere Verlässlichkeit in der Unternehmenssteuerung zu erzielen.

## Regulatorische Vorgaben zur Datenqualität

Die Verantwortung der Geschäftsführung zur Einrichtung der erforderlichen Kontrollen zur Sicherstellung der Datenkonsistenz ist nicht neu. Das Kreditwesengesetz legt fest, dass die Richtigkeit des Rechnungswesens und der Finanzberichterstattung sowie die gesetzlichen (Melde-)Anforderungen durch die Geschäftsführung zu verantworten sind. Durch immer mehr technische Lösungen und eine immer größer werdende Zahl notwendiger Reports und gesetzlicher Vorschriften hat sich diese Verantwortung im Laufe der Zeit ausgeweitet. Die Bedeutung der Datenqualität im Meldewesen sowie in der Risikosteuerung dagegen ist ein neuer Aspekt, der durch eine Vielzahl an Leitlinien, Standards und Verwaltungsvorschriften konkretisiert wird.

Der Baseler Ausschuss für Bankenaufsicht (BCBS) fordert die Institute in dem Grundsatzpapier #239 dazu auf, das Risiko-Reporting und die zugrundeliegende Datenbereitstellung zu standardisieren und zu automatisieren.

Damit ergibt sich die Notwendigkeit, die bisher weitestgehend auf verschiedene Risikoarten isoliert ausgerichteten Prozesse und Strukturen zusammenzuführen. Die Modellierung von Daten, Pflege- und Korrekturprozessen und Auswertungen müssen vor diesem Hintergrund neu aufgearbeitet werden. Daraus ergibt sich als Anforderung an die Institute, vollständige und genaue Daten zu gewährleisten sowie Verfahren und Prozesse zur Prüfung der Datenqualität zu implementieren. Zudem soll der Datenabgleich mit anderen im Institut vorhandenen Informationen und damit die Prüfung auf Plausibilität gewährleistet werden. Das Datenqualitätsmanagement (DQM) wird somit zu einer zentralen Aufgabe.

Zur Messung der Datenqualität und zur besseren Umsetzung der Datenqualitätsprüfungen hat die EZB zusätzlich den „EZB-Datenqualitätsindikator“ eingeführt. Er besteht aus den Kriterien:

- Pünktlichkeit der Datenlieferungen,
- Vollständigkeit,
- Plausibilität der Datenpunkte,
- Richtigkeit,
- Konsistenz sowie
- Stetigkeit der Meldewerte.

Der Indikator ist Teil des Moduls „Internal Governance“, das zum „Supervisory Review and Evaluation Process“ (SREP) gehört. Die Datenqualität hat somit direkten Einfluss auf die SREP-Beurteilung. Bei festgestellten Mängeln können sich daher individuelle Kapitalaufschläge zum Nachteil der Bank ergeben. Die Vorgaben gelten nicht nur für systemrelevante Institute (SIB – Systemically Important Banks). Im Sinne des Proportionalitätsprinzips müssen auch die „Less Significant Institutions“ (LSI) die Vorgaben erfüllen. Die Proportionalität bezieht sich auf den Komplexitätsgrad der Datenverarbeitungsprozesse, der IT-Architektur sowie der Geschäfte, die die Institute tätigen. Datenqualitätsmanagement ist somit eine zentrale Aufgabe der Organisation mit unmittelbarem Einfluss auf Rankings und Bewertungen des Instituts.

## Der Kreislauf des DQM verbessert permanent die Qualität

Das DQM schafft die Grundlage, um die Datenqualität regelmäßig in einem automatisierten Prozess zu prüfen (siehe Abbildung 1). Die Prüfung basiert auf Regeln, die vorab definiert werden und Auffälligkeiten ausgeben. Sie bilden den Vergleichsmaßstab für die Qualitätsmessung. Aus den Auffälligkeiten können

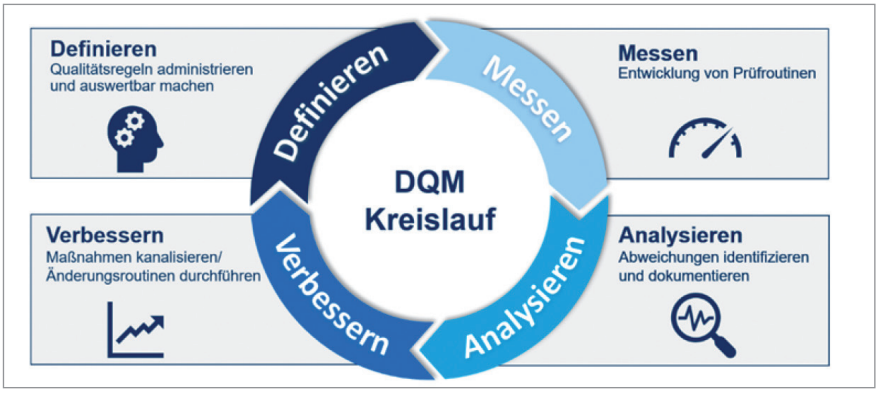



*Abbildung 1: Der DQM-Kreislauf optimiert sich kontinuierlich selbst*

wiederum neue Regeln abgeleitet werden, die erneut in die Qualitätsprüfung übergeben werden. Damit bildet sich ein Kreislauf, der sich kontinuierlich optimiert.

## Definieren – Kriterien zur Bestimmung der Datenqualität

Aus den Regularien lassen sich die für den DQM-Kreislauf nötigen Prüfkriterien ableiten. Das BCBS Dokument 239 definiert Grundsätze für die Beschaffenheit der Risikodaten und Aggregationskapazitäten. Die movisco AG hat daraus neun zentrale Kriterien zur Einschätzung der Datenqualität abgeleitet, die somit in die Messung der Datenqualität im Rahmen eines DQM eingehen.

1. ***Vollständigkeit:*** Daten gelten nur dann als vollständig, wenn sie granular auf Einzelsatzebenen vorliegen. Beim Extrahieren, Transformieren und Laden von Daten aus anderen Datenquellen (ETL-Prozess) beschreibt die Vollständigkeit, ob Inhalte und Daten vollständig übernommen worden sind.

2. ***Korrektheit:*** Vereinfacht gesagt, ist ein Datensatz dann korrekt, wenn er mit der Realität übereinstimmt. Gemessen wird die Korrektheit durch den Vergleich mit anderen Daten, deren Korrektheit bestätigt wurde. Es ist auch möglich, eine Plausibilitätsregel aufzustellen, um diese zum Vergleich heranzuziehen.

3. ***Konsistenz:*** Die Eigenschaften eines Datensatzes dürfen keine logischen Widersprüche untereinander oder zu anderen Datensätzen innerhalb einer Datenbank aufweisen. Inkonsistente Daten führen zwangsläufig zu Glaubwürdigkeitsproblemen im Reporting.

4. ***Eindeutigkeit:*** Eindeutig ist ein Datensatz, wenn die damit beschriebenen realen Objekte nur einmal dargestellt werden. Die Information ist also eindeutig, wenn sie nicht redundant ist.

5. ***Referentielle Integrität:*** Dies ist eine Besonderheit von Datenbanken. Die referentielle Integrität ist dann gegeben, wenn die Werte für Attribute eines Fremdschlüssels auch als Attribute des Primärschlüssels vorhanden sind.

6. ***Konformität:*** Daten müssen dem definierten Format entsprechen.

7. ***Gültigkeit:*** Die Daten entsprechen den definierten Wertebereichen.

8. ***Aktualität:*** Zum Zeitpunkt der Verarbeitung oder der Auswertung müssen die Daten Gültigkeit besitzen und institutsweit auf dem neuesten Stand sein. Somit bilden die Daten die tatsächliche Eigenschaft eines Objekts zeitnah ab. Die Geschäftsregeln definieren den Zeitpunkt, zu dem die Daten als aktuell angesehen werden.

9. ***Anpassungsfähigkeit:*** Die Prüfinstrumente sollen flexibel und anpassungsfähig sein, um auf neue Prüfanforderungen reagieren zu können. Durch eine Definition der Datenqualitätsmerkmale wird eine nachvollziehbare Messung der inhaltlichen Qualität der Daten erst möglich.

## Messen – Regeln messen die Datenqualität

Für die Messung der Datenqualität müssen Prüfroutinen entwickelt werden, mit deren Hilfe dann – z. B. in einer Zentralen Datenbasis – Abweichungen frühzeitig erkannt werden können. Die Regeln können in drei Gruppen zusammengefasst werden:

1. formaltechnische Regeln,
2. inhaltliche Regeln und
3. qualitative Regeln.

Formaltechnische und inhaltliche Regeln können verhältnismäßig einfach aufgestellt und beschrieben werden. Zur Vollständigkeitsprüfung eines Datensatzes kann eine Formel gebildet werden, die den Quotienten aus der Anzahl der gefüllten Attribute und der Anzahl der vorhandenen Attribute errechnet. Entspricht er nicht dem Wert 1, gibt es eine Abweichung. Die Korrektheit eines Datensatzes wird inhaltlich geprüft, indem das vorliegende Ergebnis mit Referenzwerten verglichen wird, beispielsweise mit der Bilanz oder dem Abschluss des Vormonats. Als Instrument eignet sich hier etwa der Kontenabgleich aus der Finanzbuchhaltung.

Deutlich schwieriger ist die Definition von qualitativen Regeln. Manuelle Prüfungen durch einen Fachverantwortlichen sind eine Option, die aber Ressourcen bindet und nicht skalierbar ist. Es ist somit wichtig, die Dimensionen der Datenqualität zu definieren, die besonders kritisch für Geschäftsprozesse sind.

Bei der Definition der Qualitätsregeln kann es sinnvoll sein, nicht die Gesamtmenge der vorhandenen Datensätze als Maßstab heranzuziehen. Soll beispielsweise das Feld „Geburtsdatum“ überprüft werden, ist dies nur bei natürlichen Personen sinnvoll. Fehlt der Eintrag bei einer juristischen Person, handelt es sich somit nicht um einen Fehler.

## Umgang mit Abweichungen

Sind die Regeln zur Messung der Datenqualität definiert, ist zu bestimmen, wie bei Abweichungen vorgegangen werden soll. Die Definition von Qualitätszielen und der Umgang mit Abweichungen unterliegen einer Abwägung zwischen den Geschäftsrisiken, die sich durch geringere Datengüte ergeben, und den Kosten, um das Ziel zu erreichen. Hier sind externe und interne Vorgaben zu unterscheiden (siehe Abbildung 2).

*Abbildung 2: Interne und externe Vorgaben*

Im Sinne der Erreichung einer höheren Datenqualität müssen die festgestellten Abweichungen und Fehler analysiert werden. Um die Qualität nachhaltig zu steigern, ist es notwendig, Fehlerquellen zu eliminieren. Aus der Analyse sind also entsprechende Lehren zu ziehen.

## Analysieren – Fehleranalyse zur Erreichung besserer Datenqualität

Der bisher dargestellte Prozess würde nur für eine vorübergehende Erhöhung der Datenqualität sorgen. Abweichungen wären zwar erkannt und könnten manuell bearbeitet werden; ohne allerdings die Ursachen für die Abweichungen zu ermitteln, könnten Fehler regelmäßig auftreten. Somit gilt es, sich auf die Suche nach den Ursachen zu machen, um die entstandenen Fehler aufzudecken. Im Zusammenspiel aus Daten, den Prozessen und dem Faktor Mensch ergibt sich eine Vielzahl an möglichen Ursachen für Mängel in der Datenqualität. Um die Ursachen zu ermitteln, darf der Fokus nicht ausschließlich auf den Daten selbst liegen. Technische Systeme und Plausibilitätsprüfungen leisten hier nur einen Teilbeitrag.

Genauso wichtig ist es, auch das Wissen in den Köpfen der Mitarbeiter zu nutzen. Befragungen und Diskussionen mit den Mitarbeitern führen in der Analyse oftmals zu den Fehlerquellen. Beispielsweise können Fehler bereits während der Datenerfassung auftreten. Es existieren in der Regel exakte Vorgaben, wie Informationen in das jeweilige System einzugeben sind. In einer Analyse resp. Diskussion mit den Mitarbeitern ergibt sich dann aber nicht selten das Ergebnis, dass praktische oder zeitliche Umstände dazu führen, dass diese Vorgaben gerade nicht eingehalten werden.

Auch das zugrundeliegende Datenmodell kann sich als eine Fehlerquelle herausstellen, wenn dessen Komplexität die korrekte Umsetzung erschwert. Entscheidend für die Aufarbeitung der Datenqualitätsmängel ist, die verschiedenen Fehlerursachen konsequent zu dokumentieren, um aus den daraus resultierenden Lehren Verbesserungen zu erreichen.

## Verbessern – Verbesserungen umsetzen

Zur Verbesserung der Datenqualität werden technische und nichttechnische Maßnahmen als Lehren aus der Fehleranalyse ergriffen. Am Beispiel des Attributs „Telefonnummer“ veranschaulicht, könnte eine technische Maßnahme darin bestehen, das Feld zu einem Pflichtfeld zu machen, das ausgefüllt sein muss, um die Eingabe ordnungsgemäß abzuschließen. Nichttechnische Maßnahmen könnten bspw. klare Vorgaben und Arbeitsanweisungen für die Mitarbeiter bei der Datenerfassung sein. Die in der Analyse identifizierten fehlerhaften Datensätze müssen indes bereinigt werden. Im Sinne des DQM-Kreislaufs ist dies kein einmaliger, sondern ein wiederkehrender Prozess. In der Praxis haben sich folgende fünf Methoden zur Bereinigung von Daten etabliert:

- ***Entfernung von fehlerhaften Daten:*** Nachweislich fehlerhafte Daten werden systemseitig aus dem Datenhaushalt entfernt.
- ***Ersetzung durch andere Daten:*** Fehlerhafte Datensätze können durch Referenzdatenbestände korrigiert werden. Fehlerhafte Adressinformationen können z. B. durch externe Adressdaten korrigiert werden.
- ***Ableitung aus anderen Daten:*** Fehlerhafte Daten werden aus anderen Daten abgeleitet. Eine fehlende Summe wird beispielsweise durch Addition der Einzelpositionen ermittelt.
- ***Verwendung von Default-Werten:*** Fehlerhafte Werte können auch durch das Setzen von voreingestellten Werten ersetzt werden. Das funktioniert allerdings nur dann, wenn sich für das Feld ein sinnvoller Default-Wert ermitteln lässt. Ein klassischer Lösungswert ist das Attribut „Sonstiges“.
- ***Entfernung von Dubletten:*** Dubletten beschreiben den gleichen realen Sachverhalt in mehreren Datensätzen. Die Datensätze müssen im System nicht deckungsgleich sein. Bei der Eliminierung der Dubletten muss vorher eine Konsolidierung erfolgen, welcher Datensatz die korrekten Informationen beinhaltet.

*Schritt für Schritt bessere Datenqualität mit dem DQM-Kreislauf*

## Ihre Vorteile: Datenqualitätsmanagement mit der movisco AG

Regulatorische Vorgaben und geschäftliches Interesse gleichermaßen machen eine hohe Datenqualität zu einer absoluten Notwendigkeit in jeder Bank. Wie dieser Fachbeitrag gezeigt hat, erfordert bereits die Einhaltung der Compliance vom Bankensektor, aktives Datenqualitätsmanagement zu betreiben. Hier unterstützt Sie die movisco AG mit ihrer langjährigen Expertise im Bankenumfeld wirkungsvoll. Mit der Einführung unseres Datenqualitätsmanagements optimieren Sie Ihren Datenhaushalt und schaffen mit korrekten Informationen die Basis für ein effizientes Reporting. Somit stellt sich Ihr Haus optimal auf, um die zahlreichen regulatorischen Anforderungen zu erfüllen und mit korrekten Daten und Informationen die Basis für eine solide Unternehmenssteuerung zu schaffen. Mit dem von der movisco AG implementierten DQM-Kreislauf optimieren Sie Schritt für Schritt die Qualität Ihrer Daten und erreichen die notwendige Transparenz über Ihren Datenhaushalt von der alle Geschäftsbereiche in Ihrem Institut profitieren.

***Fachbeiträge zu weiterführenden Themen können Sie auf der Website von movisco unter Research kostenlos herunterladen.***

movisco ANALYTICS

*movisco AG*
*Osterbekstraße 90a, 22083 Hamburg*
*E-Mail: info@movisco.com*
*Internet: www.movisco.com*