d-fine

The Case for Data Quality in AFC-Compliance

An often-overlooked key factor for compliance

Content

d-fine

01.   Introduction

The German regulatory framework for combating terrorist financing (CTF) and money laundering (ML) is one of the most stringent in the world, designed to safeguard the integrity of the financial system. This framework is primarily governed by the Geldwäschegesetz (GwG)[1], or German Money Laundering Act, which sets obligations for e.g. customer due diligence (CDD), risk management and risk analysis, transaction monitoring, transparency register, and reporting suspicious activities. These requirements are further strengthened by complementary regulations such as various interpretive guidelines[2,3] issued by BaFin, Germany's financial supervisory authority. Additionally, guidelines from international regulatory authorities like EBA (e.g. the risk factor guidelines[4]) and standard setters like FATF[5] or Wolfsberg Group[6] have to be taken into account.

The GwG and Kreditwesengesetz (KWG)[7] mandate financial institutions to implement robust Anti Money Laundering (AML) systems capable of identifying and mitigating ML and TF risks effectively. This includes the establishment of systems to collect, verify, and monitor customer data throughout the customer lifecycle. BaFin's AuAs BT KI[3] guidelines underscore the necessity of integrating automated systems with accurate and transparent data flows to enhance risk management and ensure compliance with regulatory expectations.

Although AML/CTF regulation often does not provide any clear requirements towards data quality, regulatory compliance obviously depends on maintaining data quality. Accurate customer records are essential for effective risk categorization, while real-time updates of customer data enable institutions to respond swiftly to emerging risks. In addition to accuracy and timeliness, the consistency of data across systems is crucial to ensure reliability and prevent discrepancies that could compromise AML efforts.

Before highlighting the importance of data quality in terms of stakeholder processes, the following subchapter provides a brief introduction to data quality.

1.1   What is data quality?

Data quality[8] refers to the degree to which data meets the needs of its users. It describes the ability of data to provide reliable, relevant, and useful information for various purposes, whether for decision-making, regulatory compliance, or system operations. High data quality is essential to ensure efficient processes and the desired value creation.

1 Gesetz über das Aufspüren von Gewinnen aus schweren Straftaten („Geldwäschegesetz,")[Online]. Available: https://www.gesetze-im-internet.de/gwg_2017/.

2 BaFin, „Auslegungs- und Anwendungshinweise zum Geldwäschegesetz," 2024. [Online]. Available: https://www.bafin.de/SharedDocs/Downloads/DE/Auslegungsentscheidung/dl_ae_auas_gw.html;jsessionid=8FC435F3C278DA8C563D691D7F18EF-BE.internet951?nn=19659504.

3 BaFin, „Auslegungs- und Anwendungshinweise zum Geldwäschegesetz – Besonderer Teil für Kreditinstitute," 2021. [Online]. Available: https://www.bafin.de/SharedDocs/Downloads/DE/Auslegungsentscheidung/dl_ae_aua_bt_ki_gw.html.

4 EBA, „Guidelines on ML/TF risk factors," 2024. [Online]. Available: https://www.eba.europa.eu/legacy/regulation-and-policy/regulatory-activities/anti-money-laundering-and-countering-financing-1.

5 FATF, „Publications," [Online]. Available: https://www.fatf-gafi.org/en/publications.html.

6 WolfsbergGroup. [Online]. Available: https://wolfsberg-group.org/.

7 Gesetz über das Kreditwesen ("KWG) [Online]. Available: https://www.gesetze-im-internet.de/kredwg/.

8 W. &. Strong, 1996. [Online]. Available: http://mitiq.mit.edu/Documents/Publications/TDQMpub/14_Beyond_Accuracy.pdf.

Data quality is evaluated across several dimensions, which are often interconnected and mutually influential. The typical dimensions of data quality[9] include:

1 **Accuracy:** Data reflects reality and contains no errors.
   Example: The correct address of a customer in a banking system matches their actual residential address, not an outdated or incorrect entry.

2 **Completeness:** All required data fields are present and fully populated.
   Example: In KYC (Know Your Customer) processes, fields such as name, date of birth, nationality, and ID number must be fully completed.

3 **Consistency:** Data is consistent and free from contradictions across systems and databases.
   Example: A customer's phone number should be identical in all systems, such as the CRM system and the AML/KYC database.

4 **Timeliness:** Data is up-to-date and reflects the most recent changes.
   Example: If a customer changes their address, this information should be promptly updated across all relevant systems to avoid communication errors or regulatory issues.

5 **Accessibility:** Data is easily accessible to authorized users without obstacles or delays.
   Example: Compliance teams must have quick access to transaction data to analyze suspicious cases in a timely manner.

6 **Relevance:** Data is suitable and useful for the specific purpose or process.
   Example: For AML monitoring, transaction data and customer risk profiles based on KYC data are more relevant than marketing preferences.

7 **Uniqueness:** Data contains no duplicates, and each data element occurs only once.
   Example: A customer should only be registered once in the database with a unique ID to avoid confusion.

8 **Traceability:** The origin and history of data are documented and can be traced.
   Example: The source of a customer's information (e.g., from an identity document or a database) should be identifiable to ensure data integrity and audit trail (who, what, when).

9 **Reliability:** Data is credible and trustworthy.
   Example: Data from an official, validated source is considered more reliable than information from unofficial or unverified sources.

10 **Validity:** A data quality dimension that refers to the conformity of data with a specific format or business rules.
   Example: Numerical information like account numbers or birthdays must be entered in the specified/correct format.

---

9 CloverDX, „8 dimensions of data quality," 2022. [Online]. Available: https://www.cloverdx.com/blog/8-dimensions-data-quality.

02.

How major AML/CTF relevant processes depend on data quality

Data quality is an important cornerstone of AML compliance, influencing the effectiveness of various processes that are crucial to identifying, managing, and mitigating risks. Among these, the organization's risk analysis framework is heavily dependent on accurate and complete data. Without reliable data, it is nearly impossible to identify vulnerabilities to ML or TF activities effectively. Institutions rely on detailed customer profiles, transaction histories etc. to assess their exposure to risks, which in turn enables the development of targeted mitigation strategies.

The following subsections highlight some key processes which are strongly impacted by data quality.

2.1 Know Your Customer (KYC) processes

Know Your Customer (KYC) processes, which includes both onboarding and ongoing due diligence, play a crucial role in mitigating ML and TF risks. When customers are onboarded, institutions must verify their identity and assess their risk level based on the information provided. High-quality data ensures that this process is not only efficient but also effective in flagging potential risks. Furthermore, ongoing due diligence requires continuous monitoring of customer data to identify any changes that might elevate risk levels. This process, which is mandated under both German and EU AML regulations, underscores the necessity of maintaining accurate and up-to-date records.

As simplified in the following diagram, the collection or recording of AML-relevant data in most cases starts with the client onboarding when a new business relationship is established.
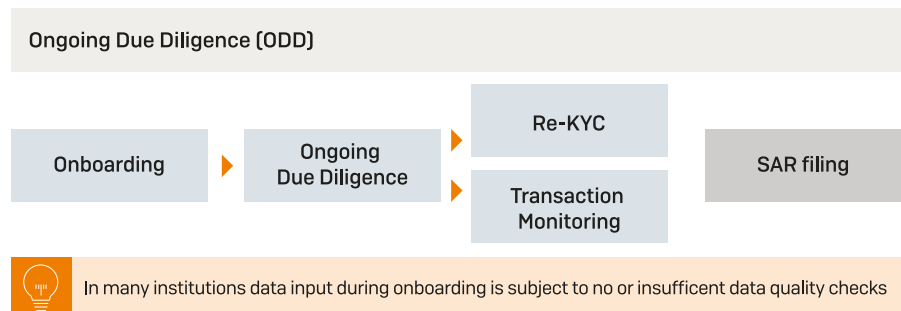
Figure 1: The basic building blocks of KYC processes require high data quality to be an effective part of anti-financial crime compliance.

d-fine

Customer risk ratings

The calculation of customer risk ratings is another area where data quality is paramount. Risk ratings guide institutions in determining the level of scrutiny required for each customer, and errors or inconsistencies in the underlying data can lead to misclassification. A misclassified customer may either be subjected to excessive scrutiny, straining resources, or insufficient oversight, exposing the institution to regulatory and reputational risks.

Accurate Data, Confident Compliance.

As observed in many institutions, weaknesses in data quality are created due to unsound processes and can built up to extensive backlogs of inconsistent or incorrect customer data. Obviously, this imposes large risks since follow-up processes can be heavily dependent in their risk-mitigating capabilities as will be discussed below.

Accurate Risk Categorization

Data quality directly impacts the precision of customer risk ratings. When customer profiles are populated with accurate and verified information—such as personal identification details, transaction histories, and behavioral patterns—risk rating models can more effectively identify high-risk individuals. Conversely, errors or inconsistencies in data may lead to misclassification, exposing organizations to significant regulatory or reputational risks. A customer incorrectly rated as low risk might avoid necessary scrutiny, while an unjustified high-risk classification could lead to lost business opportunities and strained customer relationships.
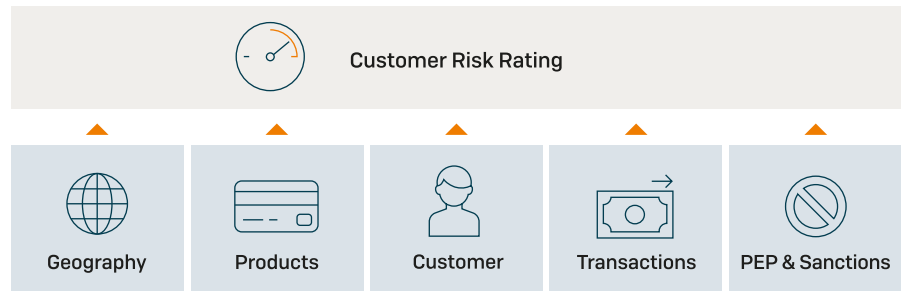

Customer Risk Rating				
Geography	Products	Customer	Transactions	PEP & Sanctions

Figure 2: High data quality is crucial for an appropriate customer risk rating.

Compliance and Regulatory Alignment

Regulatory frameworks, such as Germany's GwG and the EU's AML directives, mandate that institutions implement robust risk-based approaches. Risk ratings are central to this requirement, as they determine the level of due diligence and monitoring a customer should undergo. Poor data quality can result in inaccurate ratings, leading to failures in compliance audits or penalties for inadequate risk management. High-quality data ensures that institutions meet regulatory expectations by maintaining consistency and accuracy in their risk assessment processes.

Enhanced Monitoring and Decision-Making

Effective transaction monitoring systems often rely on customer risk ratings to prioritize suspicious activity. A well-calibrated risk rating based on reliable data helps these systems function efficiently by focusing resources on high-risk cases. Inaccurate or incomplete data can skew these priorities, causing inefficiencies such as high false-positive rates, wasted investigative resources, or overlooked legitimate risks. High-quality data enables organizations to make informed, strategic decisions about resource allocation and risk mitigation.

> ## Conclusion
>
> Data quality is not just a technical or operational concern; it is a strategic imperative that underpins the effectiveness of customer risk ratings. By ensuring data accuracy, completeness, and timeliness, organizations can achieve better compliance and operational efficiency, establishing a competitive advantage in an increasingly data-driven world.

2.3 Transaction monitoring

Transaction monitoring systems rely on accurate and comprehensive data to detect anomalies that may indicate suspicious activities. These systems use pattern recognition algorithms to analyze transaction data and flag potential risks. Poor data quality can result in a high rate of false positives, overwhelming compliance teams and diverting attention from genuine threats. Conversely, false negatives may allow suspicious activities to go undetected, potentially leading to severe regulatory and reputational consequences. The following illustration depicts a typical setup for monitoring transactions utilizing data from different silos such as customer master data, transaction data, politically exposed person (PEP) & Sanctions lists etc.
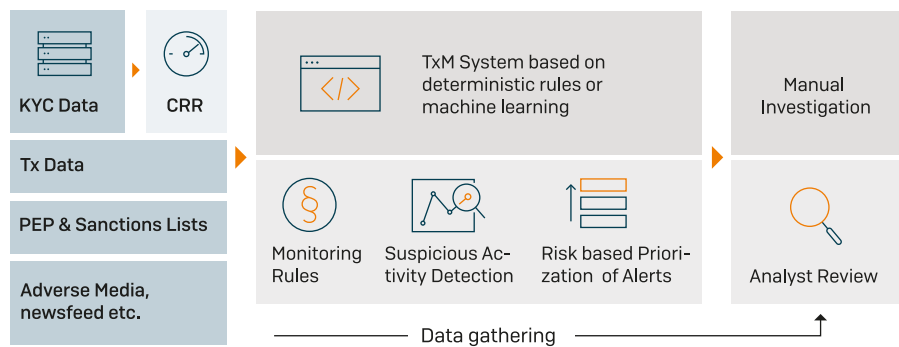

Figure 3: In transaction monitoring, very good data quality is a prerequisite for an effective and efficient monitoring process and for avoiding a high number of false positives.

The primary purpose of transaction monitoring is to identify unusual patterns or behaviors that could indicate money laundering, fraud, or other financial crimes. High-quality data enhances the accuracy of these systems by providing a complete and reliable picture of customer activities. For instance, accurate customer profiles, transaction histories, and geolocation data allow monitoring systems to distinguish between legitimate and suspicious transactions effectively. Conversely, poor-quality data can lead to missed red flags or, alternatively, unnecessary false positives that disrupt operations.

Hence it becomes clear that the ability of risk mitigation of a monitoring system is heavily dependent on the data quality of preceding systems.

Reducing False Positives and Negatives
False positives and negatives are critical issues in transaction monitoring. False positives occur when legitimate transactions are flagged as suspicious, wasting investigative resources, while potentially delaying investigation of true positives. False negatives, on the other hand, involve failing to detect genuinely suspicious activities, exposing organizations to significant regulatory and reputational risks. High-quality data reduces both by improving the precision of detection algorithms and ensuring comprehensive input for risk models.

Enabling Compliance and Regulatory Adherence
Regulatory authorities worldwide[10], including BaFin in Germany, mandate rigorous transaction monitoring practices. These requirements are designed to ensure timely and accurate detection of illicit activities and prompt filing of Suspicious Activity Reports (SARs). Inaccurate or incomplete data can result in delayed or erroneous SAR filings, leading to penalties such as financial fines or onboarding stops or even license revocation. High-quality data ensures that organizations meet regulatory standards by enabling consistent and reliable monitoring processes.

Enhancing Operational Efficiency
Data quality significantly impacts the operational efficiency of transaction monitoring systems. High-quality data allows monitoring algorithms to operate effectively and efficiently, reducing the need for manual reviews and interventions. Clean, well-structured data also facilitates seamless integration between monitoring systems and other compliance tools, such as case management or reporting platforms. This reduces operational bottlenecks and ensures that compliance teams can focus their efforts on high-priority cases.

Supporting Advanced Analytics and Machine Learning
Modern transaction monitoring increasingly relies on advanced analytics and machine learning to identify complex patterns and adapt to emerging threats. These technologies are highly dependent on the quality of their input data. High-quality data ensures that machine learningmodels are trained on accurate and representative datasets, improving their ability to detect sophisticated money laundering schemes. Poor data quality, by contrast, can introduce bias or errors into these models, undermining their effectiveness.

Adapting to Evolving Risks
As financial crime tactics evolve, transaction monitoring systems must adapt to new patterns and threats. High-quality data provides the flexibility needed to update monitoring rules and algorithms effectively. It also enables organizations to conduct historical analyses to identify trends and improve future detection capabilities by recalibration of the rule framework.

2.4 Sanction screening

Financial institutions are often struggling with high false positive rates in their sanctions screening process. Due to inconsistent and incomplete customer data, legitimate transactions are frequently flagged, leading to unnecessary compliance investigations and delays in processing.

---

[10] e.g. Financial Conduct Authority ("FCA"), „FCA fines Metro Bank £16m for financial crime failings," 2024. [Online]. Available: https://www.fca.org.uk/news/press-releases/fca-fines-metro-bank-16m-financial-crime-failings.

Typical challenges are induced by misspellings, outdated records, and missing identifiers leading to mismatches in screening. Similar as within the AML-regime, customer records are not harmonized across internal systems, causing discrepancies in screening results. Hence, we identify two major risks:

1   **High Number of False Positives:** Poor data quality results in excessive alerts, overwhelming compliance teams and reducing efficiency. This plays a crucial role in mitigating workforce shortages and efficiently allocating limited personnel resources.

2   **Regulatory Risk:** Inaccurate or incomplete data increases the likelihood of failing to detect sanctioned individuals or entities, leading to potential fines and reputational damages.

By implementing a Data Quality Management Framework, financial institutions can enhance sanctions compliance in the following ways:

1   **Standardizing and Enriching Customer Data:** Using validation tools to detect and correct inconsistencies and enrich missing information.

2   **Integrating a Single Source of Truth:** Unifying data across multiple systems to maintain consistency in screening.

3   **Automating Data Cleansing and Enrichment:** Leveraging AI and advanced analytics to detect anomalies, correct errors, and reduce manual interventions.

4   **Improving Matching Algorithms:** Ensuring that screening tools operate with high-quality data to reduce false positives and detect true risks more effectively.

Conclusion

Data quality is essential for the success of transaction monitoring systems. It enhances detection accuracy, reduces false positives and negatives, and ensures compliance with regulatory requirements. Moreover, it supports operational efficiency, advanced analytics, and stakeholder confidence. Organizations that prioritize data quality in transaction monitoring not only improve their compliance posture but also strengthen their ability to combat financial crime, safeguard their reputation, and build trust with customers and regulators. In today's complex financial landscape, investing in data quality is both a regulatory necessity and a strategic advantage.

03.

Typical issues and findings

Regulators such as BaFin frequently highlight recurring issues related to data quality that compromise AML compliance. One common finding is the presence of incomplete KYC records. Financial institutions often fail to collect all the necessary information during the onboarding process, leaving gaps in customer profiles. These gaps can have a cascading effect on other AML processes, such as risk assessment, customer risk ratings, investigation activities, transaction monitoring etc. reducing their effectiveness.

Legacy systems and data silos
Another issue often identified by regulators is the fragmentation of data across multiple systems. Many institutions operate with siloed systems that do not communicate effectively with each other, resulting in inconsistencies in customer profiles and transaction histories. These inconsistencies can lead to errors in risk assessments and delays in detecting suspicious activities, both of which are critical compliance failures.

Delays in investigation and SAR-filing
Inaccurate or delayed filing of SARs is another issue that draws regulatory scrutiny. Institutions are required to report suspicious activities promptly and accurately, but data quality issues often result in errors or omissions in these reports. Such deficiencies not only expose institutions to penalties but also undermine the effectiveness of regulatory efforts to combat ML and TF.

Every negative influence on filing of SARs to FIU can impose severe regulatory and financial fines, often accompanied with reputational damage (e.g. "naming & shaming").

Costs of data remediation
Remediating KYC data is a critical yet costly endeavor for financial institutions. Poor-quality data—whether incomplete, inaccurate, or outdated—not only hampers compliance with stringent regulatory requirements but also significantly increases operational costs. Hence, beyond regulatory fines, the costs of remediation include extensive manual reviews, re-verification of customer profiles, disruptions to business operations, and reputational damage.

Effective KYC remediation requires a comprehensive overhaul of data systems and processes, often involving investments in technology, staff training, and robust data governance frameworks. While the initial expenditure may seem high, ensuring high-quality KYC data reduces long-term costs by streamlining compliance efforts, enhancing operational efficiency, and mitigating the risk of false positives in transaction monitoring systems. Moreover, reliable KYC data supports better customer experiences, enhancing trust and customer loyalty.

This dual focus on compliance and cost efficiency highlights the strategic importance of maintaining high data quality as a foundation for sustainable AML operations.

04.

How we can help

d-fine supports clients with proven expertise in analytical, quantitative, and technological solutions while crossing the gap to regulatory requirements. From this unique position we support organizations in addressing the critical challenge of data quality. Data quality is the cornerstone of operational efficiency, regulatory compliance, and robust processes in AML/CTF or Sanctions compliance.

d-fine's holistic approach ensures that client organizations can achieve these goals effectively and efficiently.

Understanding and Assessing Data Quality Challenges
The first step in improving data quality is a comprehensive assessment of the client's current data landscape. d-fine collaborates with clients to identify gaps, inconsistencies, and inefficiencies in their data architecture. Using advanced approaches and proven methodologies, d-fine evaluates the completeness, accuracy, consistency, validity and timeliness of data across multiple systems. By aligning this analysis with the client's regulatory obligations, such as those under Germany's GwG etc., d-fine ensures that the recommendations are both practical and compliant.

Data Remediation
We have extensive experience in enhancing quality of customer master data in cases where organisations have e.g. not followed best practices, did not fulfil onboarding/ongoing due diligence obligations etc. We support our clients with custom solutions for data verification, updating and various other data quality measurements, applying proven data science methods, thereby gaining greater value from internal data records by inheritance logics or from external, trustworthy data sources.

Designing Tailored Solutions
Every organization has unique data infrastructures and data requirements depending on its industry, size, and regulatory environment. d-fine specializes in designing tailored solutions that address specific client needs while leveraging best practices. For example, in financial institutions, d-fine can design data architectures and workflows supported by data quality checks that integrate high-quality data into Know Your Customer (KYC) processes, risk scoring, and transaction monitoring systems. d-fine's ability to align technical solutions with business objectives ensures that data quality improvements directly enhance critical stakeholder processes, such as onboarding, risk analysis, and reporting.

Implementing Advanced Technology
d-fine's technological expertise allows us to implement state-of-the-art tools and platforms that automate and optimize data management. By utilizing data integration platforms, machine learning algorithms, and cloud-based systems, d-fine ensures that client organizations have access to scalable and efficient data quality solutions. For example, implementing automated validation rules and exception management processes can significantly reduce errors and ensure real-time data updates, thereby supporting compliance and operational excellence.

Summary

d-fine's proven integrated approach to data quality improvement empowers organizations to address complex challenges effectively. By combining technological innovation, industry expertise, and a deep understanding of regulatory frameworks, we enable our clients to achieve sustained improvements in data quality, driving better business outcomes while establishing compliance in an increasingly complex landscape.

Authors

Christopher Hasenberg
Manager and AFC Compliance expert
d-fine GmbH, Frankfurt
Christopher.Hasenberg@d-fine.com

Dr Ulrich Lechner
Senior Manager and AFC Compliance expert
d-fine GmbH, Frankfurt
Ulrich.Lechner@d-fine.com

d-fine

d-fine