# The Importance of Data Quality in tackling Financial Crime

*By Standard Chartered Bank, December 2018*

## Introduction

While physical cash remains an issue, most financial crime today that passes through Financial institutions involves electronic movement of funds or securities. We process well over a billion transactions a year for and on behalf of millions of customers and counterparties all over the world. The data lake that this produces is both significant and is fed from hundreds of different data sources. The importance of Data in fighting financial crime cannot be overestimated and with it therefore the importance of Data Quality. Here we share the thoughts of our Chief Data Officer on what we have learned as it relates to fighting financial crime and how Data Quality is a key component.

## Data Quality

Most global banks have had at least a decade or more of experience in managing data quality for a variety of reasons. Banks have known that capital calculations under the Basel supervisory regime can vary significantly based on the accuracy and completeness of the data used. They have also been criticised (and occasionally, fined) for poor data quality in regulatory reporting.

More recently, the Basel Committee's BCBS 239 guidelines for risk data aggregation and risk reporting have provided comprehensive guidance on ensuring data quality in systemically important Banks.

**Data Quality and why its important** - For example, screening customers and transactions against lists of sanctioned entities is heavily influenced by the data passing through those filters. Incorrect spellings of names, dummy dates of birth and incomplete or missing addresses will either result in missed sanction alerts, or lead to a large volume of false alerts that are difficult to investigate. Similarly, an AML transaction monitoring system is only as good as the data that passes through it. Not knowing that two sets of transactions relate to the same accountholder could result in incorrect conclusions. Even where a bank 'knows' their customer, not having critical data elements to drive a client risk assessment may result in underestimating the risk or overestimating the risk driving the wrong response.

Despite this, Banks have faced challenges in ensuring that Data used for detecting and preventing financial crime is fit for purpose.

This is because financial crime controls (a) are far more dependent on 'unstructured' data elements (e.g. names and addresses of customers) than other risk calculations; (b) use data that has passed through numerous hops across different parts of the Bank before reaching them; and (c) are more dependent on *external* data (i.e., data not generated or controlled by the Bank) than other types of risk controls.

Against a backdrop of data volumes that have grown over decades through many different systems, often complicated by mergers and acquisitions that have resulted in a heterogeneous legacy, this makes for an interesting challenge.


## How to Improve Data Quality

Standard Chartered has invested tens of millions of dollars in getting a better handle around the data needed for financial crime controls. Our journey is by no means complete, but here are some of the key lessons we have learnt over the years

- **Data Quality is meaningless in itself.** It only makes sense in the context in which data is used. For example, even within the context of financial crime, the name of a client is
  - irrelevant for *most anti-money-laundering algorithms*
  - somewhat relevant for *AML investigations* (though not the exact spelling)
  - extremely relevant for *sanctions compliance* (name must match CDD document)
- **"Getting data right at source" is nowhere near enough**. Data is rarely consumed at the point at which is first created. Typically, it flows through multiple bank systems/ processes before it gets used. So even if data is "correct at source", it could go wrong
  - In the hops from source to ultimate consumer (e.g. payment originator/ beneficiary information getting modified inadvertently before it leaves the Bank)
  - At the consuming end (e.g. misinterpreting 'country of CDD' as 'country of incorporation')
- **It really is cross-cutting, everyone has a role to play**. Data entry is typically done by teams in operations back offices, and in a digital world increasingly by clients themselves. Technology teams must make sure that data flows/ transformations happen as specified. But leaving it to technology and operations alone cannot ensure data quality. They cannot, for example,
  - Interpret the requirements on behalf of a data consumer (e.g., should mailing address be sent for sanctions screening?)

  - Keep the requirements up to date over time (e.g., should we start sanctions screening for domestic payments in country X?)
  - Agree whether requirements are in line with upstream business rules/ practices (e.g., is capturing mailing address for every client in line with policy/ business practice)
  - Fund or execute the clean-up of historical data quality issues (e.g., correct invalid Dates of Birth, lack of a single marker to identify all correspondent banking accounts)

- **Data quality related controls must be embedded into the bank's existing risk management framework.** Improving data quality is not a one-time exercise. It requires the implementation of controls, both preventative (e.g., maker checker controls, system validations) and detective (e.g., annual reviews of the rules used to map bank data to financial crime control systems or even better ongoing controls to detect data issues soonest). Instead of treating these as a stand-alone construct, we found that linking them to the existing risk and controls framework of the Bank was more effective and sustainable. This also enables greater ownership from relevant parties – e.g. a business leader starts viewing data quality complaints from their sanctions screening utility not as an arms-length issue, but as something that is intrinsic to their own business. That 'back-to-front' linkage can help create the necessary prioritisation and focus for data quality.

- **Now is a great time to move the needle on data quality.** The internal argument in Banks for better data quality has traditionally been around regulatory compliance and the risk of fines. But there is a much more potent, and positive, story to be told around data quality today. With the use of data, analytics and Artificial Intelligence (AI) to drive front-line business outcomes, businesses have become much more aware of the need to ensure good data quality. For example, better data quality allows businesses to ensure
  - Customer experience for digital natives (not having to update the same address twice)
  - More effective customer analytics (knowing the full set of suppliers and buyers of a client, predicting product needs)
  - Better lending decisions (flexing credit limits based on ongoing transaction history)

## A "Data Quality Management Framework"

At SCB, our Data Quality Management Framework (DQMF) embodies the principles outlined above. The Framework focuses on manual data entry and the points where data is transferred downstream from upstream systems or processes, for ultimate consumption in risk management and reporting systems.

- **Aspirations:** the DQMF provides a common language across the Bank, it defines roles and responsibilities of key stakeholders, and it outlines the governance regarding strategy setting, oversight and escalations of issues.
- **Enablers:** across Technology, Process and Education, the DQMF supports data quality by providing a bank-wide data repository, uniform reference data, and tools to manage data taxonomy, data lineage and data profiling.
- **Outcomes:** evidence of an effective DQMF comes from improved client experience, increased productivity and end-to-end digitation in operational processes, greater use of data and analytics in processes and strategy setting, overall fewer operational issues related to data, timely identification, escalations and resolution of new issues, and bank-wide awareness of the importance of data and data quality.

Within SCB's risk management framework, the Chief Data Officer (CDO) is the owner of the DQMF and has delegated 2nd line of defence oversight over data quality and the impact of that on processes.

## New Technology Can Help

Advances in AI and data analytics are also making seemingly intractable data quality issues easier to manage. For example, natural language processing on unstructured text can enable more accurate and timely extraction of relevant data from client documentation and transaction messages. Investing in a data lake makes it easier to bring relevant data into one place (subject to regulatory constraints).

Powerful, user-friendly data search, preparation and visualisation tools make it easier for end users to explore and analyse data. Machine Learning can be leveraged to detect and even self-correct some types of data quality issues.

Last, where available some government and regulatory initiatives can also be leveraged, for example to integrate client onboarding systems with government ID databases.

However, rarely will the tools fully fix the data issues, they will mostly only help to manage them.

***"Data quality should not be an after thought, but instead is a fundamental part of any Bank's strategy, including financial crime risk management."***

**Shameek Kundu - Chief Data Officer, Standard Chartered Bank**

## Conclusion

Managing data quality is critical to the detection and prevention of financial crime outcomes. It has always been a challenge, and will remain so in the foreseeable future.

But with a robust framework to manage the risk stemming from data quality issues, the advances in data analytics and AI and the broader business imperative to manage data as an asset, Banks have a real opportunity to make a significant difference in this space.

.