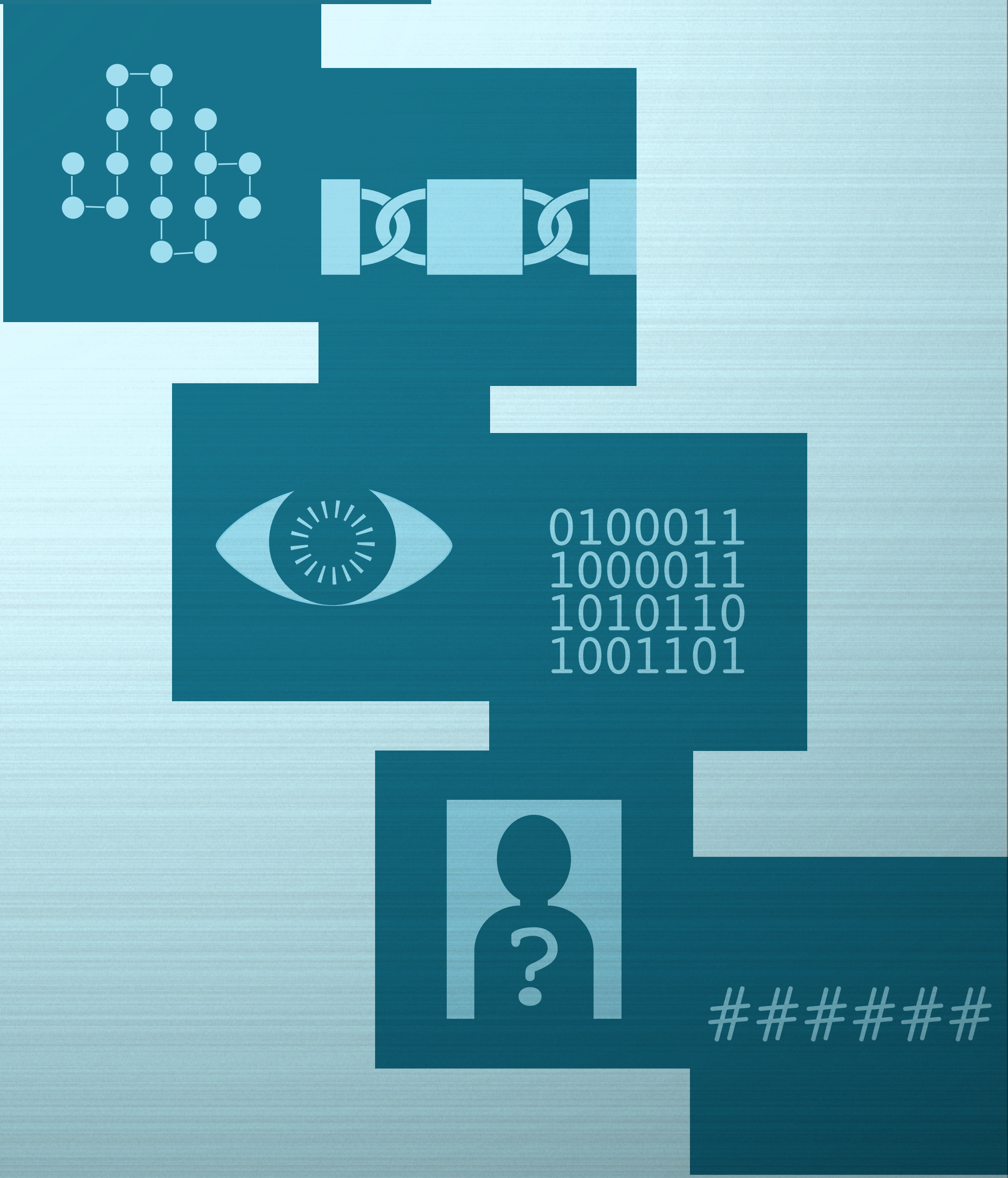
Center for Global Development

Fixing AML

Can New Technology Help Address the De-risking Dilemma?

JIM WOODSOME
VIJAYA RAMACHANDRAN



Center for Global Development
2055 L Street, NW
Fifth Floor
Washington, DC 20036

www.cgdev.org

CONTENTS

ACKNOWLEDGEMENTS

We are grateful to the experts and practitioners who shared their knowledge with us. In particular, we would like to thank Bart Claeys, Tarannum Islam, Paul Janssens, David McLaughlin, Simon Moss, Kevin Petrasic, Meral Ruesing, Benjamin Saul, Gurjeet Singh, David Stewart, and Stephan Wolf.

In addition, we would like to thank Christopher Calabia, Amy Jerrett-Kotsenas, Abed Mutemi, Rosita Najmi, and Michael Pisa for their valuable feedback.

CGD is grateful for contributions from the Bill & Melinda Gates Foundation and the Lakeshore Foundation in support of this work.

Any errors or omissions remain the responsibility of the authors.

ACRONYMS

ABA	American Bankers Association
AI	artificial intelligence
AML	anti-money laundering
API	application programming interface
BAFT	Bankers Association for Finance and Trade
BCBS	Basel Committee on Banking Supervision
BIC	business identifier code
CBR	correspondent banking relationship
CDD	customer due diligence
CFT	countering the financing of terrorism
CGD	Center for Global Development
CIP	customer identification program
CPMI	Committee on Payments and Market Infrastructures
DFID	Department for International Development (United Kingdom)
DLT	distributed ledger technology
DTCC	Depository Trust & Clearing Corporation
ELT	extract, load, and transform
ETL	extract, transform, and load
FATF	Financial Action Task Force
FCA	Financial Conduct Authority (United Kingdom)
FinCEN	Financial Crimes Enforcement Network (US Department of the Treasury bureau)
FSB	Financial Stability Board
GLEIF	Global Legal Entity Identifier Foundation
GLEIS	Global Legal Entity Identifier System
IIB	Institute of International Bankers
IIF	Institute of International Finance
IMF	International Monetary Fund

ISO	International Organization for Standardization
KYC	know your customer
KYCC	know your customer’s customer
LEI	legal entity identifier
LEI ROC	Legal Entity Identifier Regulatory Oversight Committee
LOU	Local Operating Unit (also referred to as LEI Operating Unit)
MAS	Monetary Authority of Singapore
MTO	money transfer operator
NPO	nonprofit organization
P2P	peer-to-peer
PEP	politically exposed person
PIN	personal identification number
PMPG	Payments Market Practice Group
SAR	suspicious activity report
SWIFT	Society for Worldwide Interbank Financial Telecommunication

EXECUTIVE SUMMARY

In 2015, a Center for Global Development (CGD) working group on the unintended consequences of anti-money laundering (AML) and countering the financing of terrorism (CFT) policies argued that the policies that have been put in place to counter financial crimes may also have unintentional and costly consequences for people in poor countries. The report's authors identified the problem of "de-risking." In other words, AML/CFT policies have had a chilling effect on banks' willingness to undertake cross-border transactions. Banks are unwilling to do business in markets perceived to be risky (or low in transaction volume) in part because of the perceived high cost of compliance.

Who are the losers from de-risking? The analysis of the CGD working group points to the families of migrant workers; small businesses that need to access working capital or trade finance; and recipients of lifesaving aid in active conflict, post-conflict, or post-disaster situations. And sometimes, AML/CFT policies may be self-defeating to the extent that they reduce the transparency of financial flows.

One consequence of de-risking may be the decline in correspondent banking services, which are critical to cross-border financial transactions. Worldwide, the number of correspondent banking relationships has declined by more than 6 percent since 2011. Small and fragile countries have been especially affected.

Even while policy solutions to address de-risking are being implemented, new technologies have emerged to address de-risking by increasing the efficiency and effectiveness of AML/CFT compliance by financial institutions. These new technologies may enhance transparency and information-sharing capabilities, facilitate automation and interoperability between platforms and institutions, mutualize certain compliance functions, and improve banks' ability to accurately identify illicit activity.

This report, to the best of our knowledge, is the first comprehensive effort to assess six key new technologies and their potential to solve the de-risking problem. These include *know-your-customer (KYC) utilities, big data, machine learning, distributed ledger technology (DLT), legal entity identifiers (LEIs),* and *biometrics*. These new technologies (in use and on the horizon) may make it easier to conduct AML/CFT compliance, which in turn might tip banks' cost-benefit calculation and make holding correspondent banking accounts with clients in poor countries more likely.

With a view to educating policymakers, regulators, and the broader audience interested in addressing the de-risking problem, we describe what these technologies are and how they work. We examine what parts of the AML/CFT compliance workflow they can improve, including customer identification and verification, customer due diligence, and transaction monitoring. We also examine the limitations of these technologies and the barriers to adoption they face. Finally, we offer recommendations for how policymakers and regulators can responsibly support the adoption of these technologies.

KYC utilities are central repositories for customer due diligence (CDD) information. By centralizing information collection and verification, KYC utilities can reduce the amount of information that has to be exchanged bilaterally between correspondent banks and their respondents, thereby reducing the time banks spend conducting CDD investigations. KYC utilities may also help facilitate the adoption of a baseline dataset for CDD information. Several KYC utilities were launched in 2014 and 2015, catering to different client segments, including the correspondent banking sector. In one business model, the information provider (in the case of correspondent banking, the respondent bank) uploads its information for free, and the information consumer (the correspondent bank) pays a fee to access the information. Although KYC utilities emerged organically in response to market demand, most industry bodies have urged regulators to produce more explicit guidance as to how much banks may rely on these services, in order to further increase utilization.

Big data refers to datasets that are high in volume, high in velocity, and high in variety, and therefore require systems and analytical techniques that differ from those used for traditional datasets. Compared with relational databases, big data applications offer more scalable storage capacity and processing. They also allow many different types of data to be stored in one place, so compliance staff spend less time gathering information from disparate sources. Most important, they can greatly expand the range and scope of information available for KYC and suspicious transaction investigations. Big data applications are typically paired with advanced analytics engines—including machine learning programs—that can help identify complex patterns and relationships in the data that might have otherwise gone undetected.

Machine learning is a type of artificial intelligence—itself a branch of computer science—that allows computers to improve their performance at a task through repeated iterations. There are three broad types of machine learning—supervised, unsupervised, and reinforcement learning. With supervised learning, the machine learning program analyzes a dataset to build a model that best predicts a predefined output. In contrast, with unsupervised learning, the machine learning program is not given a predefined output—rather, it explores the data on its own, looking for patterns and relationships in the dataset. Reinforcement learning falls between the other two, with the algorithm receiving general feedback on its performance, but without a specific predefined output to aim for. Machine learning may be used to augment or transform a number of compliance functions, including those for developing more sophisticated customer typologies and for more accurately monitoring transactions. These uses could simultaneously cut down on false alerts and identify new or hitherto undetected illicit finance techniques. Banks may benefit from more leeway to explore these new technologies. Banks would also benefit from more government feedback on the suspicious activity reports (SARs) they file, which would help them to further hone their detection capabilities.

DLT is a way of securely organizing data on a peer-to-peer network of computers. In a blockchain, which is a type of DLT, data modifications, such as transactions, are recorded in time-stamped blocks. Each block is connected to previous blocks, forming a chain. Modifications are confirmed and stored by all users on the network, which makes the ledger difficult to tamper with. Although blockchain technology is most commonly associated with virtual currencies, such as Bitcoin, the basic technology has a number of other potential use cases, including uses in regulatory compliance. In particular, DLT may be used for securely storing and sharing KYC information, as well as for cheaper and more secure international payments. This technology is yet to be widely adopted, but single-use cases are emerging in different parts of the world.

LEIs are unique alphanumeric identifiers, like barcodes, that connect to reference datasets held in a public database. Any legal entity that makes financial transactions or enters into contracts may request an LEI. In many countries, especially developed ones, LEIs are increasingly mandated by regulation. To date, more than 1 million LEIs have been issued worldwide. By serving as common identifiers, LEIs can enable different platforms, organizational units, and institutions to refer to entities clearly and without any ambiguity. This interoperability can, in turn, facilitate greater automation and information sharing. In addition, the reference datasets can serve as a starting point for CDD. A further extension of the LEI would be to include it in payment messages to identify originators and beneficiaries, which would further enhance the transparency of international payments. However, this would require changes to payment message formats and to banks' IT systems, as well as more widespread adoption of the LEI outside of the financial sector and also in developing countries.

Biometrics use distinctive physiological or behavioral characteristics to authenticate a person's identity and control his or her access to a system. Natural persons are not eligible for LEIs except in limited circumstances, so a separate standard is needed for identifying individuals. Biometrics are more robust than other authentication factors, such as passwords and tokens, as they are generally more secure and easier to use. Biometrics are being used to address the "identification gap" that exists in many developing countries. This use, in turn, could make it easier for banks to conduct customer identification, verification, and due diligence, which may bolster the confidence of their correspondent banks. However, most biometric identification systems are being developed at the national level, meaning that individual identification is still fragmented at the international level, hindering the use of biometrics for international payments. Work is required to develop an internationally recognized and interoperable identification system for natural persons.

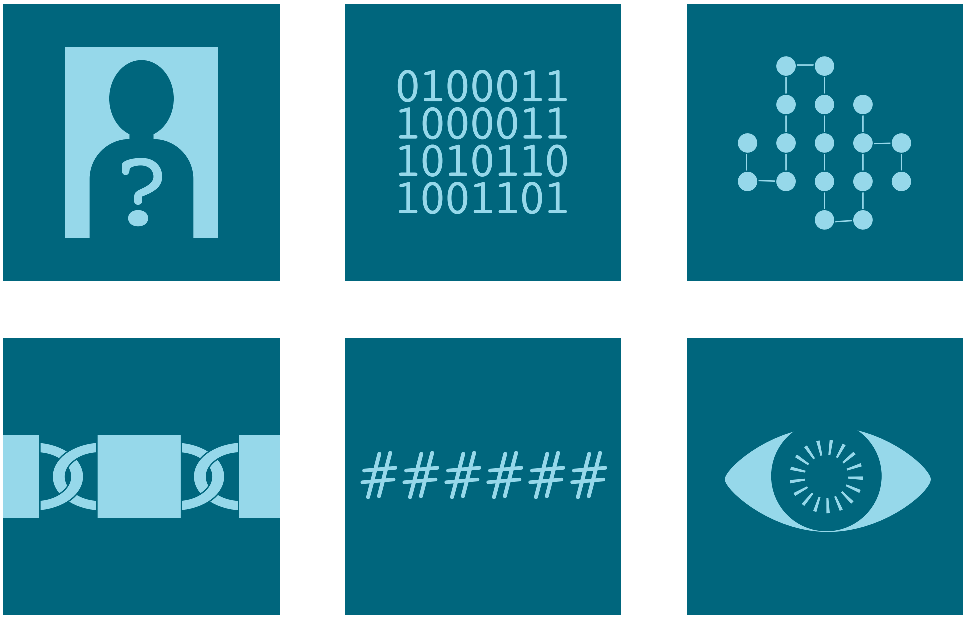
Table 1 provides a summary of the recommendations for stakeholders to responsibly move each of the above-described technologies forward. We discuss these recommendations in more detail in the report.

Table 1. Summary of Policy Recommendations

Organizations Involved	Recommendation
KYC Utilities	
National regulators	■ Give further consideration as to whether and to what degree financial institutions can rely on third parties for customer identification and due diligence, and offer further guidance, if necessary. It is important that banks understand the degree to which they can rely on KYC utilities or other third-party information sharing mechanisms. ■ Provide clarity on who bears (or is allowed to bear) liability if CDD information is incorrect. ■ Consider whether to establish regulatory regimes for regulating and monitoring KYC utilities.
Standard-setting bodies and international organizations	■ Explore steps necessary to establish KYC utilities that cater to money transfer operators (MTOs) and nonprofit organizations (NPOs). As part of this exploration, work with banks, MTOs, and NPOs to develop standardized due diligence questionnaires, similar to the Wolfsberg Group's Correspondent Banking Due Diligence Questionnaire. Also consider whether economic support is needed to make such solutions a reality, or whether they can be market driven. ■ Continue to engage on developing issues related to KYC utilities. ■ Explore whether it is possible for third parties also to conduct risk assessments themselves, as opposed to simply providing information for risk assessments.
Big Data	
National regulators and international organizations	■ Determine whether local privacy and data sharing laws pose a challenge to the integration of these datasets and whether these laws can or should be amended without compromising privacy.
Artificial Intelligence/Machine Learning	
National regulators	■ Share feedback on SAR submissions. ■ Allow financial institutions to share data, so as to expand the pool of information that machine learning programs can learn from. ■ Consider a regulatory sandbox to allow financial institutions to experiment with machine learning solutions.
Distributed Ledger Technology/Blockchain	
National regulators	■ Consider how to amend data sharing and privacy laws to enable sharing of identification, due diligence, and transaction data between banks or between banks and authorities.
DLT participants	■ Begin working to establish common interoperability standards, to ensure that different arrangements can integrate with each other.[a]
Legal Entity Identifiers	
Standard-setting bodies	■ Determine whether LEIs can be used for customer identification and verification and for due diligence, and provide relevant guidance.
National regulators in countries affected by de-risking	■ Look for ways to promote LEI issuance. ■ Improve business registries and other relevant information sources that Local Operating Units use to validate information.
Financial institutions	■ Help customers obtain LEIs, especially in countries affected by de-risking.
International Organization for Standardization (ISO)	■ Continue work on how best to incorporate the LEI into the new payments messaging format.
Biometrics	
Standard-setting bodies	■ Explore what steps are needed to develop an internationally recognized, interoperable digital identification system for natural persons.
National regulators	■ Continue to develop biometric-based national ID systems with robust privacy controls.

[a] CPMI, 2017, p. 18.

1 BACKGROUND: BANKS' AML/CFT OBLIGATIONS, DE-RISKING, AND NEW TECHNOLOGICAL SOLUTIONS FOR COMPLIANCE



The de-risking phenomenon

In recent years, many large international banks have partially or wholly withdrawn from certain countries and classes of customer. They have done so for numerous reasons, one of which is that they find it increasingly difficult and expensive to serve such customers while complying with laws designed to curb illicit finance. As a result, the affected parties may find it harder to access certain financial services, particularly for cross-border payments.

This phenomenon, commonly referred to as de-risking, poses a challenge to financial inclusion and, more broadly, to international financial integration, economic growth, and poverty reduction. Often, the populations most affected by de-risking are already vulnerable, such as migrant workers and people living in fragile countries. In addition, de-risking undermines transparency by pushing legitimate transactions into unregulated channels, making it harder to detect illicit financial flows.

The de-risking of correspondent banks

Among the sectors affected by de-risking is correspondent banking. Correspondent banking is a type of bilateral interbank relationship in which one bank (the correspondent bank) provides financial services to another bank (the respondent bank) and usually, by extension, to the respondent bank's customers.[1] Large international banks act as correspondents to local and regional banks, which depend on these relationships for access to foreign markets and currencies.[2] A decline in correspondent banking relationships (CBRs) can make it harder for export-oriented businesses to secure trade finance, for migrants to send remittances back home, and for nonprofit organizations (NPOs) to provide humanitarian aid abroad.[3] In the most dire scenarios, a total or near-total loss of CBRs could effectively cut a country off from the international financial system.

In recent years, numerous studies have documented a worldwide decline in the number of CBRs. In a 2015 World Bank survey, half of national banking authorities, more than half of local and regional banks, and three-quarters of large international banks all reported declines in international CBRs during the period 2012–2015.[4] Similarly, a recent analysis of data from the Society for Worldwide Interbank Financial Telecommunication (SWIFT) by the Financial Stability Board (FSB) found that the number of active CBRs worldwide had declined by 6 percent from 2011 through the end of 2016.[5] This decline is a global phenomenon, affecting virtually all regions and major international currencies.[6] Importantly, the number

[1] "Correspondent clearing" is a correspondent banking service whereby the correspondent bank clears payments on behalf of the respondent bank's customers.

[2] World Bank, 2017a, p. 105.

[3] CGD Working Group, 2015, pp. vii-viii.

[4] Ninety-one banking authorities completed the survey, as did 20 large international banks and 170 local and regional banks. See World Bank, 2015b, pp. 5, 12.

[5] FSB, 2017a, p. 1.

[6] FSB, 2017a, p. 2.

of correspondent banking accounts that transact in either US dollars or euros—the two most important international currencies—have each declined by 15 percent.[7] The FSB also found evidence that correspondent banking networks were becoming more concentrated, which, it warned, might make them more fragile as well as lower competition and raise the cost of international payments.[8]

Some countries and regions have been affected worse than others. Among the regions that have been especially affected are the Caribbean, the Pacific Islands, the Middle East and North Africa, central Asia, and sub-Saharan Africa.[9] The loss of CBRs in places like the Pacific Islands is especially worrisome, given that these countries typically rely on just a handful of CBRs to begin with.[10]

The countries that have experienced the steepest declines in correspondent banking activity have tended to be small or else perceived as being at high risk of money laundering, terrorist financing, or sanctions evasion.[11] High-risk countries include those under US or multilateral sanctions, such as Iran and North Korea. They also include countries afflicted by high levels of crime, terrorism, war, or civil unrest, such as Syria, Yemen, Libya, and Venezuela.[12] Countries with large offshore banking sectors, such as some in the Caribbean, have experienced significant declines as well.[13] Small countries face a different problem. Correspondent banking is typically a fee-based service, and small countries may not be able to generate a sufficient volume of payments to cover the costs of servicing them, especially when compliance costs are rising.[14]

The de-risking of money transfer operators and NPOs

Money transfer operators (MTOs) and NPOs have also been affected by de-risking. MTOs facilitate the flow of remittances, especially to populations not well served by banks. NPOs are often critical to the delivery of humanitarian aid. The de-risking of MTOs and NPOs is often bound up in the de-risking of CBRs, but it can also occur independently.[15]

The de-risking of MTOs is widespread. In another 2015 World Bank survey, 28 percent of MTOs reported having lost access to traditional financial services.[16] Of these, a quarter were unable to find work-around solutions.[17] In addition, 45 percent

[7] FSB, 2017a, p. 14.

[8] FSB, 2017a, p. 33.

[9] IMF, 2017, p. 14; World Bank, 2015b, p. 40.

[10] Alwazir et al., 2017, p. 22.

[11] FSB, 2017a, p. 4; IMF, 2017, pp. 12–13.

[12] FSB, 2017a, pp. 22–23.

[13] Alleyne et al., 2017, p. 13.

[14] IMF, 2017, pp. 21–22.

[15] MTOs and NPOs may lose financial access as a direct result of their banks' correspondent accounts being terminated. Banks may also close MTOs' and NPOs' accounts, or otherwise restrict their activity, in order to appease their correspondent banks. In other cases, however, the de-risking of MTOs and NPOs is unrelated to developments in their banks' CBRs.

[16] Eighty-two MTOs responded to the survey. See World Bank, 2015a, p. 1.

[17] Reported work-around solutions include "a) using other MTOs, b) operating via cash management companies and physically transporting cash, and c) using personal bank accounts" (World Bank, 2015a, p. 19).

of MTOs reported that their agents had lost access to banking services. This is of particular concern because MTOs rely on agents to reach poor and rural areas they would otherwise not operate in. The World Bank warned of diminished financial access in these areas, particularly for remittance recipients.[18]

The de-risking of NPOs is less studied, but available evidence confirms that humanitarian organizations have been affected. A survey of internationally active US-based NPOs reported widespread financial access problems, including account closures (reported by 6 percent of surveyed NPOs) and account denials (10 percent). More frequently, they included less severe but still disruptive financial access problems, such as wire transfer delays and unusual documentation requests. Altogether, approximately two-thirds of respondents reported encountering some type of financial access problem, and 15 percent reported encountering these problems regularly or constantly.[19]

AML/CFT compliance as a driver of de-risking

De-risking has many drivers, two of which are rising compliance costs and falling risk tolerances. In a recent survey, the FSB found that 20 percent of CBR terminations had to do with anti-money laundering (AML), countering the financing of terrorism (CFT), and sanctions. Issues included rising compliance costs as well as a lack of confidence in the respondent banks' risk controls.[20]

National governments charge financial institutions with significant responsibilities for countering illicit finance. These responsibilities include performing due diligence on their customers, monitoring accounts and transactions for suspicious activity, and reporting suspicious activities to their governments (see the appendix to this chapter for more details).

In recent years, regulators in the United States, the United Kingdom, and several other countries have raised their AML/CFT compliance standards. At the same time, they have cracked down on financial institutions that have fallen short of these standards.

In response, banks have attempted to scale up their existing compliance procedures or enact new ones. In some cases, the procedures did not scale well or else did not work as intended, leading to widening inefficiencies and spiraling costs.[21]

Surveys indicate that AML/CFT compliance costs are rising. KPMG's 2014 annual AML survey reported that 78 percent of survey respondents had increased their spending on AML compliance since 2011, with a fifth of survey respondents claiming spending increases of more than 50 percent.[22] A LexisNexis Risk Solutions survey of financial institutions (mostly banks) in six Asian countries found that

[18] World Bank, 2015a, pp. 19–20.

[19] Eckert, Guinane, and Hall, 2017, pp. vi and 38–41.

[20] FSB, 2017a, p. 43.

[21] PA Consulting Group, 2017, p. 13.

[22] The main areas of investment were transaction monitoring, know-your-customer capabilities, and staff. See KPMG, 2014, pp. 13–14.

supporting correspondent banking was among the top drivers of AML/CFT investments, and the top driver for 11 percent of respondents.[23]

For some banks, compliance costs have risen so much that the provision of correspondent banking services is no longer profitable for them, and there "is no business justification for continuing to engage in correspondent banking."[24]

The potential of regulatory technology

In the face of de-risking, public- and private-sector stakeholders have sought ways to lower the compliance burden without lowering standards. Policymakers, regulators, standard-setting bodies, and industry associations now regard the mitigation of de-risking as a major imperative. A number of important efforts are underway to assess the scale and scope of the de-risking problem, as well as to devise regulatory and technological solutions to it.

Regulatory technology (regtech) may offer a partial solution to de-risking. *Regtech* refers to "the use of new technologies to solve regulatory and compliance requirements more effectively and efficiently."[25] It may be applied to a number of different regulatory compliance functions, including AML/CFT.

Regtech may help alleviate de-risking by lowering compliance costs and by improving risk management capabilities. In particular, regtech innovations may improve banks' abilities to conduct due diligence on their direct customers and to monitor transactions more effectively. Certain innovations even have the potential to shed new light on the risk profiles of payment originators and beneficiaries, bringing new transparency to correspondent banking. Such innovations could enable banks to more confidently service higher-risk or lower-revenue customers, while at the same time improving their ability to detect illicit finance. All of these results, in turn, could reduce banks' incentive to de-risk.[26]

Financial institutions are optimistic about regtech. In a recent survey of AML compliance staff, almost 60 percent of respondents reported that regtech had improved their ability to handle AML/CFT, know-your-customer (KYC), and sanctions compliance. More than half expected to increase their investments in regtech over the next three to five years.[27]

Purpose and structure of the report

The purpose of this report is to examine regtech solutions that might help alleviate de-risking. Because de-risking is largely due to decisions made by large international banks in developed countries, the report focuses on regtech solutions that are being (or can be) adopted by those banks.

[23] LexisNexis Risk Solutions, 2016, p. 4.

[24] CPMI, 2016, p. 12.

[25] IIF, 2016, p. 3.

[26] IIF, 2017a, p. 2.

[27] Dow Jones and SWIFT, 2017, p. 36.

Six innovations, in various stages of development and adoption, show particular promise. These include KYC utilities for the sharing of customer due diligence (CDD) information; big data systems for the management of large, heterogeneous datasets and enhanced monitoring; machine learning for advanced pattern recognition and prediction; blockchain technology for more secure identification and information sharing; legal entity identifiers (LEIs) for the identification of legal entities; and biometric identifiers for the identification of natural persons.

For consistency and ease of reading, the technology profiles follow the same basic structure. Each profile begins with a description of the technology (what it is and how it works), followed by an examination of its advantages and use cases, as well as its challenges and limitations. Each profile concludes with a discussion of the technology's prospects for adoption and a set of recommendations for regulators, policymakers, and standard-setting bodies.

Appendix:
Financial Institutions' AML/CFT Compliance Responsibilities and Workflow

In order to analyze how technology can improve banks' AML/CFT compliance processes, it is necessary first to understand banks' compliance workflow. This appendix provides a general overview of banks' AML/CFT compliance responsibilities as set out by the Basel Committee on Banking Supervision (BCBS) and the Financial Action Task Force (FATF), two international standard-setting bodies.[28] *It also gives some attention to the particular compliance challenges posed by correspondent banking.*

Under the current international AML/CFT regulatory framework, financial institutions have significant responsibility for monitoring and countering illicit finance. These responsibilities fall into three broad, overarching categories:

- Customer identification, verification, and due diligence procedures (collectively referred to as "know-your-customer," or KYC)
- Ongoing monitoring of accounts and transactions for suspicious activity
- Reporting suspected illicit financial activities to the government

Risk management principles are established by international standard-setting bodies, including the FATF and the BCBS. Laws and regulations are set at the national level. Further guidance is provided by industry groups.[29]

There are general principles and laws, and specific guidance for particular services or activities, such as correspondent banking. Correspondent banking activities—especially international correspondent clearing—are considered especially vulnerable to illicit finance abuse owing to their lack of transparency and moral hazard issues (see Box 1 for more).

[28] The appendix does not delve into national regulations.

[29] For correspondent banking, the two industry groups that provide guidance on AML/CFT compliance are the Wolfsberg Group and The Clearing House.

Box 1. Why Correspondent Banking Is Considered Vulnerable to Illicit Finance Abuse

Correspondent banking is the provision of financial services to other banks, either foreign or domestic. Correspondent clearing is a type of correspondent banking service in which the correspondent bank clears payments on behalf of the respondent's customers.

Correspondent banking is considered particularly vulnerable to money laundering and terrorist financing owing to its lack of transparency and moral hazard problems. As the Wolfsberg Group lays out, "a correspondent bank is effectively acting as [the respondent bank's] agent or conduit, executing and/or processing payments or other transactions for the [respondent bank's] customers."[a] Further, "the [correspondent bank] may have no direct relationship with the underlying parties to any transaction routed through it and, in such cases, may not be in a position to verify identity or to understand fully the nature of the specific transaction, particularly when processing electronic payments (wire transfers) or clearing checks."[b]

Correspondent banking's susceptibility to money laundering was first highlighted in a 2001 US Senate report. The report, the product of a yearlong investigation, was authored by the Democratic staff of the US Senate Permanent Subcommittee on Investigations, led by Senator Carl Levin.[c]

Correspondent clearing is considered the riskiest correspondent banking activity. This is because the correspondent bank has a limited view of the respondent banks' customers. In effect, the correspondent bank must largely rely on the respondent banks' own AML/CFT controls.[d] This is especially true of international correspondent clearing, in which the respondent banks operate under different laws and may be subject to different supervision standards than the correspondent banks.

Correspondent clearing becomes even riskier when it involves nested relationships.[e] Nested relationships involve the respondent bank's customers, which may include other banks or MTOs, who have customers of their own. Nesting is not illegal, but correspondent banks must always conduct enhanced due diligence in CBRs that involve nesting.

Correspondent banking transactions are particularly challenging to monitor effectively because of their volume, speed, and fungibility.[f] As The Clearing House notes, "[these] very attributes [that] are so instrumental to the successful functioning of the payment system . . . also create vulnerability to money laundering." The Clearing House further argues that unless the payment originator and beneficiary are clearly and correctly identified in the payment message, and unless they are flagged as "problematic" in advance, it is "virtually impossible [for banks] to identify and intercept [suspicious] payments."[g] Further, once illicit funds have been introduced into the formal payments system, "it is very difficult, and in many cases impossible, to identify those funds as they move from bank to bank." [h] If the beneficiary and/or originator information is intentionally misrepresented, it is very hard for the correspondent bank to detect the misrepresentation.[i]

[a] Wolfsberg Group, 2014a, p. 1.
[b] Wolfsberg Group, 2014b, p. 2.
[c] Minority Staff of the Permanent Subcommittee on Investigations, 2001.
[d] Protiviti, 2017, p. 5.
[e] Protiviti, 2017, p. 5.
[f] The Clearing House, 2016, p. 4.
[g] The Clearing House, 2016, pp. 4-5.
[h] The Clearing House, 2016, pp. 4-5.
[i] The Clearing House, 2016, p. 5.

Appropriate programs, policies, and procedures

Financial institutions must first understand the illicit-finance threat environment they operate in and then develop appropriate programs, policies, and procedures to counter these threats. A financial institution's risk assessment should be based on the prevalence and nature of the illicit finance risk inherent in (1) the customers the bank does business with, (2) the jurisdictions it does business in, (3) the products and services it sells, and finally, (4) the distribution channels it uses.[30]

Customer identification, verification, and due diligence procedures (KYC)

Banks must establish robust customer identification, verification, and due diligence procedures. When accepting and onboarding new customers, banks must be able to positively identify them. Banks must also collect information on their new customers in order to assess their risk of engaging in illicit finance. In the case of correspondent banking, banks must also vet the respondent bank's parent company (if the respondent bank is an affiliate, subsidiary, or branch).[31]

First, banks must have a customer identification program (CIP) in place. The CIP defines "clear, systemic procedures and policies to identify and verify its customers and, where applicable, any person acting on their behalf and beneficial owner(s) of transactions."[32] The CIP also specifies the minimum set of information the bank should collect from prospective respondent banks, such as name, address, and ID number.[33]

Second, banks should verify their customers' identities using "reliable, independently sourced documents, data, or information."[34] Banks may use *documentary verification methods* (e.g., certified articles of incorporation or business licenses) as well as *non-documentary methods* (e.g., comparing the information provided by the customer with publicly available information, or checking with other banks that have worked with the customer before).[35]

Third, banks must gather sufficient information about a client to assess that client's illicit finance risk. Depending on the type of customer, this could include, for example, sources of income, country of origin, and business activities.[36]

For correspondent banking, the due diligence process should take into account three categories of risks. As outlined by the BCBS, these are (1) the risks involved in the services being provided, (2) the risks inherent to the respondent bank itself, and (3) the risks associated with the jurisdiction in which the respondent bank operates.[37]

[30] BCBS, 2017, p. 3.

[31] Wolfsberg Group, 2014a, p. 2.

[32] BCBS, 2017, p. 8. In the United States, banks have until May 11, 2018, to implement new procedures to identify beneficial owners; see Miller and Rosen, 2017, p. 9.

[33] The Clearing House, 2016, p. 8.

[34] BCBS, 2017, p. 8.

[35] The Clearing House, 2016, p. 8.

[36] BCBS, 2017, p. 9.

[37] BCBS, 2017, pp. 24–25.

The KYC process must be conducted not only during the acceptance and onboarding process, but also periodically throughout the course of the relationship. Banks must periodically review and update their CDD information. They must also rescreen their customers against lists of sanctions and politically exposed persons (PEPs) as those lists are updated.

The bank is expected to continually refine its understanding of its customers, their typology, and their behavior.[38] Only by understanding what constitutes normal behavior for a particular customer (or type of customer) can the bank identify suspicious activity.[39]

Due to its inherently risky nature, correspondent banking is never eligible for reduced or simplified CDD. Regulatory guidelines now permit reduced or simplified due diligence for certain categories of low-risk customers, but correspondent banking is not one of them.

Especially risky CBRs must be subjected to enhanced due diligence procedures. Some CBRs are considered riskier than others. These include relationships with respondent banks that involve PEPs, either as customers, as employees, or as directors. They also include relationships with respondent banks that have nested relationships with other banks or MTOs. Nested relationships are legal but must be subjected to extra scrutiny.[40]

Ongoing monitoring of accounts and transactions for suspicious activity

Banks must have transaction monitoring systems for detecting abnormal or suspicious transactions.[41] According to the BCBS, "using CDD information, a bank should be able to identify transactions that do not appear to make economic sense . . . or that are not consistent with the customer's normal and expected transactions."[42] Banks should also keep current on what known money laundering schemes are being used in their jurisdictions, customer bases, products, and services.

In the case of correspondent banking, respondent banks are responsible for making sure their payments messages are accurate and complete. It is especially important that payment messages include correct information on the payment originators and beneficiaries, so that their correspondent banks can monitor transactions for suspicious activity and comply with sanctions screening.[43]

As a general rule, correspondent banks are not expected to conduct due diligence on their respondent banks' customers (a practice referred to as "know your customer's customer," or KYCC). However, they are expected to look for unusual or suspicious patterns in the transactions of their respondent banks' customers.[44]

[38] BCBS, 2017, p. 8.

[39] BCBS, 2017, p. 8.

[40] Wolfsberg Group, 2014a, p. 6.

[41] BCBS, 2017, p. 10.

[42] BCBS, 2017, p. 11.

[43] BCBS, 2017, p. 28.

[44] KYCC has been a source of regulatory uncertainty for correspondent banks in recent years and a contributor to de-risking in the sector. International standard-setting bodies and national regulators have attempted to clarify their expectations. See, for example, FATF, 2016a, p. 4.

When the correspondent bank detects a transaction that appears unusual or suspicious, it can halt the transaction and transmit a request for information to the respondent bank.[45] Based on the respondent bank's response to the request, if the correspondent bank is satisfied that the transaction is legitimate, it can unfreeze it.

Record keeping and reporting

Banks should maintain records of all the KYC information they collect. They should make sure these records are kept up-to-date.[46]

Banks should report suspicious activities to the authorities. These authorities could be bank supervisors or law enforcement. In the United States, banks are required to report suspicious activity to the Department of the Treasury through the submission of suspicious activity reports.[47]

If banks are uncomfortable with their ability to manage the risk of a particular CBR, the BCBS encourages them to first explore alternatives to termination. Such alternatives include "limiting the services provided, real-time monitoring, sample testing of transactions, or on-site visits."[48]

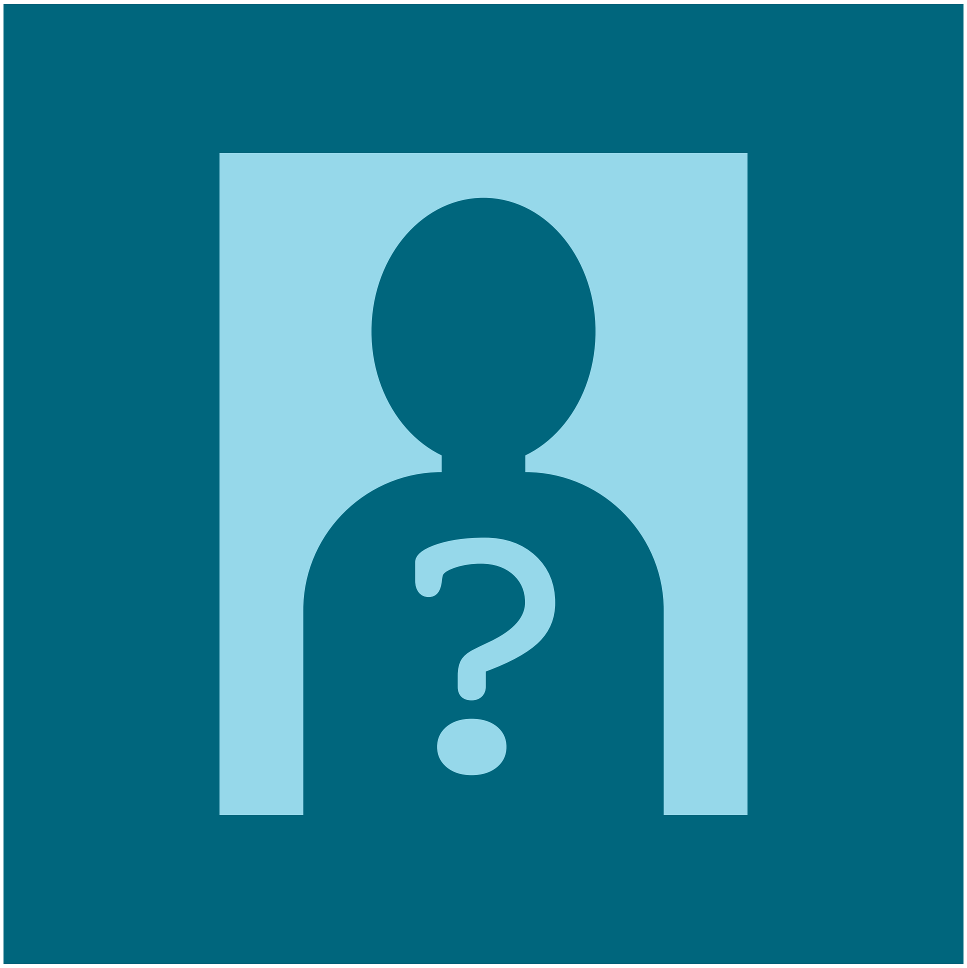
[45] BCBS, 2017, p. 29.

[46] BCBS, 2017, p. 11.

[47] Miller and Rosen, 2017, p. 7.

[48] BCBS, 2017, p. 29.

2 KYC UTILITIES



Key points

- Know-your-customer (KYC) utilities are central repositories for customer due diligence (CDD) information on financial institutions' direct clients or counterparties.
- KYC utilities may reduce industrywide compliance costs by consolidating the exchange of CDD information, thereby reducing duplicative processes.
- KYC utilities may serve as a catalyst for the development of common due diligence standards.
- A number of KYC utilities were launched in 2014 and 2015, often with the backing of one or more major financial institutions. They vary by ownership structure, business model, and target market. The Society for Worldwide Interbank Financial Telecommunication's (SWIFT's) KYC Registry is the most prominent KYC utility in the correspondent-banking space.
- KYC utilities facilitate KYC, not know-your-customer's-customer (KYCC). However, SWIFT's KYC Registry now enables KYCC analysis down to the counterparty level, though not to the account level.
- KYC utilities emerged organically in response to market demand, which demonstrates that their existence does not require government backing. However, most industry bodies argue that KYC utilities would be more fully utilized if financial regulators were to grant banks permission to rely on and support the use of KYC utilities.
- The Basel Committee on Banking Supervision (BCBS) has provided guidance on the use of KYC utilities, but so far, national regulators in the United States and elsewhere have not publicly addressed the issue.

Box 2. The 2015 Center for Global Development Report on KYC Utilities

In its 2015 report, the Center for Global Development Working Group on the Unintended Consequences of Anti–Money Laundering Policies noted that correspondent banks' desire to reduce compliance costs and regulatory risks is a driver of de-risking. The report featured SWIFT's then-year-old KYC Registry as "geared toward facilitating data sharing and making the KYCC concept less expensive and more manageable over time."[a] This, it argued, "effectively turns KYC information into what economists refer to as a *club good*, where members can share information without paying any additional costs beyond the price of admission."[b]

In the report's recommendations, the working group advocated for the more widespread use of KYC utilities:

> **Recommendation 5: Facilitate Identification and Lower the Costs of Compliance:** Better messaging standards and KYC documentation repositories. Banks and other financial institutions should redouble their efforts, with encouragement from the FSB [Financial Stability Board] and national regulators, to develop and adopt better messaging standards and implement KYC documentation repositories. . . . A Know Your Correspondent registry could reduce compliance costs by holding KYC documents from clients in a format that can be queried by banks rather than requiring customers to submit new documents that must be verified.[c]

a. CGD Working Group, 2015, p. 32.
b. CGD Working Group, 2015, p. 29.
c. CGD Working Group, 2015, p. 55.

What are KYC utilities?

KYC utilities are central repositories for CDD information on financial institutions' direct customers or counterparties.[49] They take in, cross-check, and store data and documents relevant to the conduct of customer identification, verification, and due diligence. Member financial institutions can access this information and use it to perform their own KYC checks and risk assessments.[50]

KYC utilities represent an effort on the part of banks to mutualize certain anti–money laundering (AML) and countering the financing of terrorism (CFT) compliance costs.[51] This allows member institutions to save on non-revenue-generating activities that confer little or no competitive advantage.[52]

KYC utilities are a type of shared utility. At their heart, shared utilities serve as "a single source of data or [a single point of] operation" for multiple members or customers.[53] There are also shared utilities for communications, trading and execution, clearance and settlement, cash and collateral management, and asset custody, to name just a few functional areas.[54] In the past few years, more than 40 financial-sector utilities have been launched.[55]

Shared utilities allow banks to outsource certain functions. Since the global financial crisis, banks have intensified their outsourcing of back- and middle-office activities.[56] Shared utilities are a viable outsourcing solution in functional areas in which the demand for customization is low and the potential for standardization is high.[57]

A number of KYC utilities were launched in 2014 and 2015. These included the Depository Trust & Clearing Corporation's (DTCC's) Clarient Entity Hub (since acquired by Thomson Reuters); Markit/Genpact's KYC Services; SWIFT's KYC Registry; and Thomson Reuters' Accelus Org ID (now simply Org ID) (see Table 2).

These KYC utilities were launched in direct response to banks' requests for such a service. As the cost and complexity of KYC compliance became more difficult to manage, banks began to realize that doing this on their own no longer made sense.[58]

In many cases, these utilities were owned or backed by one or more major financial institutions.[59] For example, the DTCC's Clarient Entity Hub was backed by Barclays, Bank of New York Mellon, Goldman Sachs, JPMorgan Chase, and State Street Corporation, among others. SWIFT's KYC Registry was launched in participation with Bank of America, Barclays, Citigroup, Commerzbank, Deutsche Bank,

[49] CPMI, 2016, p. 19.

[50] PwC, 2015, p. 1.

[51] Ray, 2015, slide 7.

[52] PwC, 2015, p. 3.

[53] Twiggs, 2015, p. 40.

[54] Nelson et al., 2016, p. 17.

[55] Nelson et al., 2016, p. 6.

[56] Ray, 2015, slides 4–6.

[57] Ray, 2015, slides 6–9.

[58] Interview with Bart Claeys, head of KYC Compliance Services, SWIFT, September 8, 2017.

[59] Ray, 2015, slide 6.

Table 2. Major KYC Utilities Launched in 2014/2015[a]

Service Provider	DTCC	Markit/Genpact	SWIFT	Thomson Reuters
Name of Utility	Clarient Entity Hub	KYC Services (KYC.com)	The KYC Registry	Org ID (prev. Accelus)
Year Launched	2015	2014	2014 (Dec.)	2014
Current Status	Acquired by Thomson Reuters; now part of Org ID	Operational	Operational	Operational
Information Providers (at Outset)	Investment managers Hedge funds Nonfinancial corporations	Investment managers Hedge funds Nonfinancial corporations	Respondent banks (both SWIFT and non-SWIFT users)	Investment managers Hedge funds Nonfinancial corporations
Information Consumers (at Outset)	Banks Broker/dealers	Banks Broker/dealers	Correspondent banks	Banks Broker/dealers

a. Ray, 2015, slide 19; Todd and Hochstein, 2014.

Erste Group, HSBC, ING, JPMorgan, RBI, Standard Chartered, and UniCredit.[60] SWIFT itself is a member-owned cooperative.

KYC utilities vary by ownership structure. These ownership structures include the following:

- *Intrabank utility:* An internal utility for information sharing within a single financial group or holding company.[61]
- *Industry collaboration:* A joint venture developed by two or more financial institutions.[62]
- *Utility service provider:* A market utility set up by an independent third-party vendor.[63]
- *Jurisdictional utility or national utility:* A utility focused on entities located in a particular jurisdiction (see Box 3).

KYC utilities also vary by their target market. Different KYC utilities exist to provide information on asset managers, hedge funds, and nonfinancial corporations. KYC utilities' target markets can shift over time as the utilities expand or focus their offerings.

SWIFT's KYC Registry, launched in December 2014, focuses on the correspondent banking sector. As of mid-2017, the KYC Registry had signed up more than 4,500 financial institutions in more than 200 jurisdictions, including 60 central banks.[64]

No single KYC utility dominates across all market segments. Among the various functional areas described above, "the KYC space has seen the highest number of utility offerings emerge," according to Arin Ray, an analyst for Celent, a consultancy.[65] The Committee on Payments and Market Infrastructures (CPMI) predicts

[60] Todd and Hochstein, 2014, p. 1; interview with Bart Claeys, September 8, 2017.

[61] IMF, 2017, p. 33.

[62] PwC, 2015, p. 1.

[63] PwC, 2015, p. 1.

[64] Interview with Bart Claeys, September 8, 2017.

[65] Ray, 2015, slide 11.

Box 3. Jurisdictional KYC Utilities and e-KYC

Jurisdictional KYC Utilities

A few jurisdictions are establishing their own KYC utilities. Jurisdictional KYC utilities may cover all entities, or a specific subset of entities, located in a particular jurisdiction. They may be set up as public utilities or as public-private partnerships, or they may be wholly private-sector initiatives. Countries that are in the process of establishing KYC utilities, or exploring the possibility of doing so, include Singapore and South Africa.

In South Africa, Thomson Reuters partnered with Barclays Africa, Rand Merchant Bank, Standard Bank of South Africa, and Standard Chartered to set up a KYC utility dedicated to these banks' South African corporate and institutional clients.[a]

Singapore's central bank, the Monetary Authority of Singapore (MAS), is working with a number of foreign and local banks to develop a pilot KYC utility for individuals. The project will utilize MyInfo, a platform developed by Singapore's Ministry of Finance and its Government Technology Agency, which stores verified personal information, linked to individuals' national ID numbers. MAS expects to begin piloting the utility in 2018.[b]

e-KYC

Some countries are enabling banks to connect to their national identification systems. This allows banks to access information on natural persons that has already been collected and verified by the government.

In India, banks may use the Aadhaar biometric ID system to verify their customers' identities.[c] Bank Negara Malaysia, Malaysia's central bank, has also laid the regulatory groundwork for e-KYC, with a particular view toward improving the KYC process for remittance recipients.[d]

[a] Thomson Reuters, 2016, 2017.

[b] *Finextra*, 2017.

[c] *The Economist*, 2016.

[d] Aruna, 2017.

that it is "unlikely that any single utility will emerge catering to all segments and use cases."[66] The existence of multiple KYC utilities may limit the economies of scale that can be achieved, though it may also preserve competition.[67] It is also possible that KYC utilities will ultimately settle on focusing on particular market segments or jurisdictions. In doing so, they would achieve monopoly or quasi-monopoly status within their particular domains.

[66] CPMI, 2016, p. 22.

[67] Loffi, 2016, slide 8; Ray, 2015, slide 24.

How KYC utilities work

The core service that KYC utilities provide is the collection and management of CDD information. This includes collecting the relevant data and documentation, as well as cross-checking it, organizing it, storing it, periodically updating it, and providing access to it. Risk assessment remains the purview of the member financial institutions.

KYC utilities may vary in their operating models and fee structures. In the case of SWIFT's KYC Registry, the information provider (e.g., the respondent bank) uploads its information to the KYC utility for free. Staff at the KYC Registry then review this information and cross-check it. The respondent bank controls who has access to its information. If a respondent bank is not on a utility, the correspondent bank may invite it to join as part of the onboarding process. The information consumer (e.g., the correspondent bank) accesses this information for a fee.[68] The correspondent bank may access the information through a password-protected web portal or through an application programming interface (API).

SWIFT's KYC Registry collects a baseline dataset from respondent banks comprising five categories of information. These data requirements were initially determined by a working group of 12 major correspondent banks. They are periodically updated to reflect changes in regulatory requirements and customer needs (see Table 3 for the most recent version, as of the end of 2017). In October 2017, SWIFT announced that it had aligned its data requirements with the new Wolfsberg Correspondent Banking Due Diligence Questionnaire.[69]

SWIFT's KYC Registry uses the business identifier code (BIC) as its reference number. It has begun to incorporate the legal entity identifier (LEI) where available.

Some KYC utilities offer additional services. These may include screening for politically exposed persons (PEPs), sanctions screening, negative news alerts, and certain types of analysis.

SWIFT's KYC Registry now enables KYCC analysis down to the counterparty level, but not to the account level. The SWIFT Traffic Profile provides correspondent banks with an aggregated view of their payments and trade finance transactions to and from high-risk jurisdictions (defined as jurisdictions under US or EU sanctions, as well as jurisdictions deemed high risk or noncompliant by the Financial Action Task Force, or FATF). SWIFT monitors subscribing customers' payments messages to produce the Traffic Profile data. A correspondent bank can view its exposure to risk on an aggregated basis (that is, for all sanctioned or high-risk counterparty jurisdictions), as well as disaggregated by country or counterparty (i.e., respondent bank). The SWIFT Traffic Profile does not provide the granularity to conduct KYCC on the originators and beneficiaries of transactions.[70]

[68] The exact fee structure varies from utility to utility. For example, Markit/Genpact charges an annual base licensing fee plus a variable "per entity" fee. See Ray, 2015, slides 14, 15, and 19.

[69] SWIFT, 2017.

[70] Interview with Bart Claeys, September 8, 2017; email correspondence with Bart Claeys, November 27, 2017.

Table 3. SWIFT KYC Registry Baseline Dataset[a]

Category	Information and Supporting Documentation
I. Identification of the customer	Legal names, addresses, legal forms, industry classification, registration information, and regulatory/supervision information
	Supporting documentation includes proof of regulation, business licenses, extract from registers, certification of incorporation, and certificate of change of name.
II. Ownership and management structure	Form of organization, including listing information, bearer shares information, and ownership structure with identifying data of shareholding companies (owning 10 percent or more) including trusts and foundations, and ultimate beneficial owners (owning 10 percent or more)
	Supporting data and documentation includes identifying data of key controllers (board of directors, senior executive management, supervisory board) and related supporting documents, including annual reports and key financial data.
III. Type of business and client base	Type of products and services, customer base, geographical presence and operations, and business with sanctioned countries
IV. Compliance	Compliance contacts (money laundering reporting officer, chief compliance officer for AML); responses to AML questionnaire based on the Wolfsberg Correspondent Banking Due Diligence Questionnaire 2017; copy of AML procedures; and USA PATRIOT Act, Markets in Financial Instruments Directive, and International Securities Services Association questionnaires
V. Tax information	Tax identification number; Foreign Account Tax Compliance Act information, including Global Intermediary Identification Number, contact, and form; and Common Reporting Standard information, including contact and self-certification

a. SWIFT, n.d.–a.; email correspondence with Bart Claeys, head of KYC Compliance Services, SWIFT, November 27, 2017.

The advantages of KYC utilities

KYC utilities can substantially reduce industrywide compliance costs by consolidating the exchange of CDD information. Before the advent of KYC utilities, CDD was conducted on a wholly bilateral basis, which led to widespread duplication of effort in gathering and verifying CDD information.[71] Correspondent banks had to collect information from each respondent bank they did business with, first during the onboarding process and then again periodically throughout the course of the relationship. Respondent banks, in turn, had to provide the same (or similar) information to each of their correspondents, with variations due to the lack of standardization in banks' requirements and regulatory policies.[72] A KYC utility

[71] Commenting on this state of affairs, the American Bankers Association, an industry group, wrote: "There is no good explanation for why each financial institution must replicate the effort. It is almost as though each time someone wanted to use a bridge to cross a river, the rules said that even though that's a nice bridge, well-constructed and properly maintained, you must build your own bridge for the crossing" (ABA, 2017, p. 7).

[72] This process led to "a massive exchange of documents," according to the CPMI. "The 7,000 banks that use the SWIFT network for correspondent banking have more than 1 million individual relationships, so the number of documents exchanged is presumably much higher" (CPMI, 2016, p. 19).

centralizes this process. Rather than having to provide the same information over and over again, respondent banks provide their core information to only one organization—the KYC utility. At the same time, correspondent banks can spend less time gathering and verifying information on their respondent banks if there is a core set of information in the KYC utility that they can rely on.

KYC utilities can also help banks reduce customer onboarding times. Currently, the onboarding process can take weeks or even months.[73] According to a survey conducted by SWIFT, members of its KYC Registry spent an average of 45 percent less time on due diligence than before they joined.[74]

KYC utilities may serve as a catalyst for the development of common due diligence standards. Correspondent banks have different information requirements. This lack of standardization makes it harder for respondent banks to fulfill information requests, and it increases the scope for error. Respondent banks may find it easier to ensure their information is correct if they have just one core dataset to maintain, as opposed to many.[75]

Challenges and limitations

The main limitation that KYC utilities face is a lack of clear regulatory backing at the national level. Correspondent banks are still ultimately responsible for due diligence, and it is unclear how much they can rely on information collected by KYC utilities. In the absence of assurances from regulators, supervisors, and law enforcement, banks will not be able to wholly give up their duplicative processes.[76] Moreover, the Institute of International Finance (IIF) and the Bankers Association for Finance and Trade (BAFT) have argued that, to the extent that banks are now expected to conduct due diligence on KYC utilities as well, their workload will be increased, not decreased.[77]

A number of industry groups have argued that this lack of regulatory backing limits the usefulness of KYC utilities. The IIF and BAFT have noted that in the absence of regulatory assurances, "banks have sometimes concluded they should treat [KYC utility] data as essentially the same as receiving data from the client itself, and therefore subject to an obligation to identify and verify the information."[78] In a joint letter to the BCBS, The Clearing House and the Institute of International Bankers (IIB) argued that "these tools are unlikely to materially reduce the cost of KYC compliance for correspondent banks absent a change in regulatory expectations."[79] The BCBS has made a first step toward clarifying the issue (more on that below), but many uncertainties remain at the national level.

It is unclear whether KYC utilities and financial institutions agree on what, specifically, regulatory backing should look like. The IIF recommends that KYC utilities be allowed to assume legal liability for the information they provide, which,

[73] Ray, 2015, slide 12.

[74] SWIFT, 2016.

[75] That said, most stakeholders acknowledge that even a well-designed data template will not suffice in all cases.

[76] IIF, 2016, p. 10.

[77] IIF and BAFT, 2017, p. 11.

[78] IIF and BAFT, 2015, p. 9.

[79] The Clearing House and IIB, 2017

Table 4. Summary: Advantages and Challenges/Limitations of KYC Utilities

Advantages	Challenges
■ Can reduce duplication of effort in the exchange of CDD information for both information collectors and information providers, reducing time and cost of customer onboarding and improving overall industry efficiency ■ Can lead to more reliable, higher-quality data, since information providers need focus on providing and maintaining only a single core dataset ■ Can serve as a catalyst for the creation of common due diligence standards	■ KYC utilities focus on facilitating KYC, not KYCC. SWIFT's KYC Registry now enables KYCC analysis down to the counterparty level, but currently not to the transaction level. ■ Banks are still responsible for risk assessment. ■ Banks still retain liability for accuracy of CDD. ■ The utilities do not necessarily meet all of banks' information needs. ■ Putting all the information in one place may make utilities tempting to hackers. ■ Banks need to vet KYC utilities and conduct periodic checks on data quality.

the industry group argues, is necessary to free financial institutions from having to conduct their own data checks.[80] It is unclear whether KYC utilities wish to assume such liability.

A second limitation is the lack of standardization across utilities, with respect to both data templates and refresh rates. Currently, KYC utilities tailor their data templates to the legal requirements of the jurisdictions they operate in, as well as to feedback from member institutions.[81] The problem of standardization is made more difficult by the fact that different jurisdictions have different requirements for identification and due diligence.

A third limitation is that privacy laws and bank secrecy laws may prohibit using, transferring, and storing certain categories of data, especially across borders. This can prohibit institutions from sharing important information, even internally.[82] Respondent banks may also be reluctant to share information with a KYC utility. Clients are liable for safeguarding the data they download from KYC utilities.

A fourth limitation is that the utilities do not cover all organizations and do not collect all information that a correspondent bank might want. Some information requirements will always be particular to individual correspondent banking relationships (CBRs).[83] This is particularly true when it comes to enhanced due diligence. This means that correspondent banks will always retain some responsibility for collecting and verifying CDD information.

[80] IIF, 2017a, p. 5.

[81] CPMI, 2016, p. 21.

[82] IIF, 2016, p. 10.

[83] Interview with Bart Claeys, September 8, 2017.

What regulators, policymakers, and standard-setting bodies are doing to facilitate the adoption of KYC utilities

International policymakers and standard-setting bodies have shown that they are increasingly aware of KYC utilities and are generally favorable toward them. National policymakers and regulators have generally been more circumspect, at least publicly.

The BCBS has provided detailed guidance on the use of KYC utilities, including the extent to which correspondent banks can rely on them.[84] The guidance states that "supervisors see in principle no objection to the use of [KYC] utilities in correspondent banking risk assessment processes," provided that certain conditions are met and that the correspondent bank understands it retains ultimate responsibility for CDD.[85] On the subject of reliance, the guidance explains that the level of risk will determine whether the correspondent bank needs to independently verify or augment the information it receives from the KYC utility.[86] Finally, the guidance provides a list of factors banks should consider when assessing the validity of information provided by the utility.[87]

Work is underway to define a standard baseline dataset that KYC utilities should collect. The IIF and BAFT have argued that standardization would reduce the need for correspondent banks to conduct due diligence on the KYC utilities themselves. In their view, standardization should cover "recognized international standards for data quality, the type and amount of information required, database maintenance and upkeep, audit, and governance."[88] Adherence to such standards would be carried out either by the authorities themselves or through an officially recognized certification process.

The Wolfsberg Group recently updated its Correspondent Banking Due Diligence Questionnaire, which could serve as a template for the baseline dataset.[89] The questionnaire provides a standardized set of questions for correspondent banks to ask prospective respondent-bank clients. Prior to its release, the Financial Stability Board (FSB) had stated that this updated questionnaire could be used to the define the baseline dataset that KYC utilities collect from respondent banks.[90] Subsequent to the questionnaire's release, the Wolfsberg Group's members—which include 13 global banks—announced that they would be adopting the questionnaire

[84] According to the Financial Stability Board, this fulfills the G20's commitment to clarify international regulatory expectations with respect to KYC utilities. See FSB, 2017b, p. 2.

[85] BCBS, 2017, p. 27.

[86] BCBS, 2017, p. 34.

[87] These include whether the information is sourced, when it was last updated, whether and when the utility verified the information with the source, and whether the information the utility provides is reliable, as judged by periodic data checks by the bank. See BCBS, 2017, p. 35.

[88] IIF and BAFT, 2017, p. 11.

[89] Interview with Bart Claeys, September 8, 2017.

[90] FSB, 2017b, p. 2.

as their new standard.[91] In October 2017, SWIFT announced that it had aligned its information baseline with the new questionnaire as well.[92]

The FSB is considering whether to advocate for the creation of a new International Organization for Standardization (ISO) standard for KYC utilities' baseline dataset. Another possibility is that the FSB will choose to encourage KYC utilities to adopt an existing standard, such as ISO 20022, the universal financial industry messaging scheme.[93]

US authorities have not yet publicly commented on KYC utilities. The American Bankers Association (ABA) has recently recommended that the US Department of the Treasury's Financial Crimes Enforcement Network (FinCEN) issue guidance validating the use of third parties for customer identification and due diligence, including "clear parameters" for when banks may rely on CDD information provided by third parties. It has further "urg[ed] Treasury to support [the efforts of KYC utilities] and help eliminate resistance on the part of banking regulators."[94]

Prospects for adoption

A number of KYC utilities have been established in recent years in response to market demand. This indicates that the KYC utilities' core business model—information collection and checking—can endure even in the absence of explicit regulatory support.

Uptake of KYC utility services by financial institutions is already significant. A 2016 survey by Capco found that 45 percent of respondents were using or were planning to use a utility for KYC and AML; another 18 percent said they would consider it.[95] Today, in correspondent banking, approximately 65 percent of SWIFT's users are now members of the KYC Registry.[96]

However, regulatory reforms will probably be needed for KYC utilities to achieve their full potential. Deepening and extending that business model, which has the potential to reduce systemwide compliance costs more drastically, will likely require regulatory reforms.

[91] Wolfsberg, 2017.

[92] SWIFT, 2017

[93] FSB, 2017b, p. 17.

[94] ABA, 2017, pp. 6–7.

[95] Capco and Finextra, 2016, p. 12.

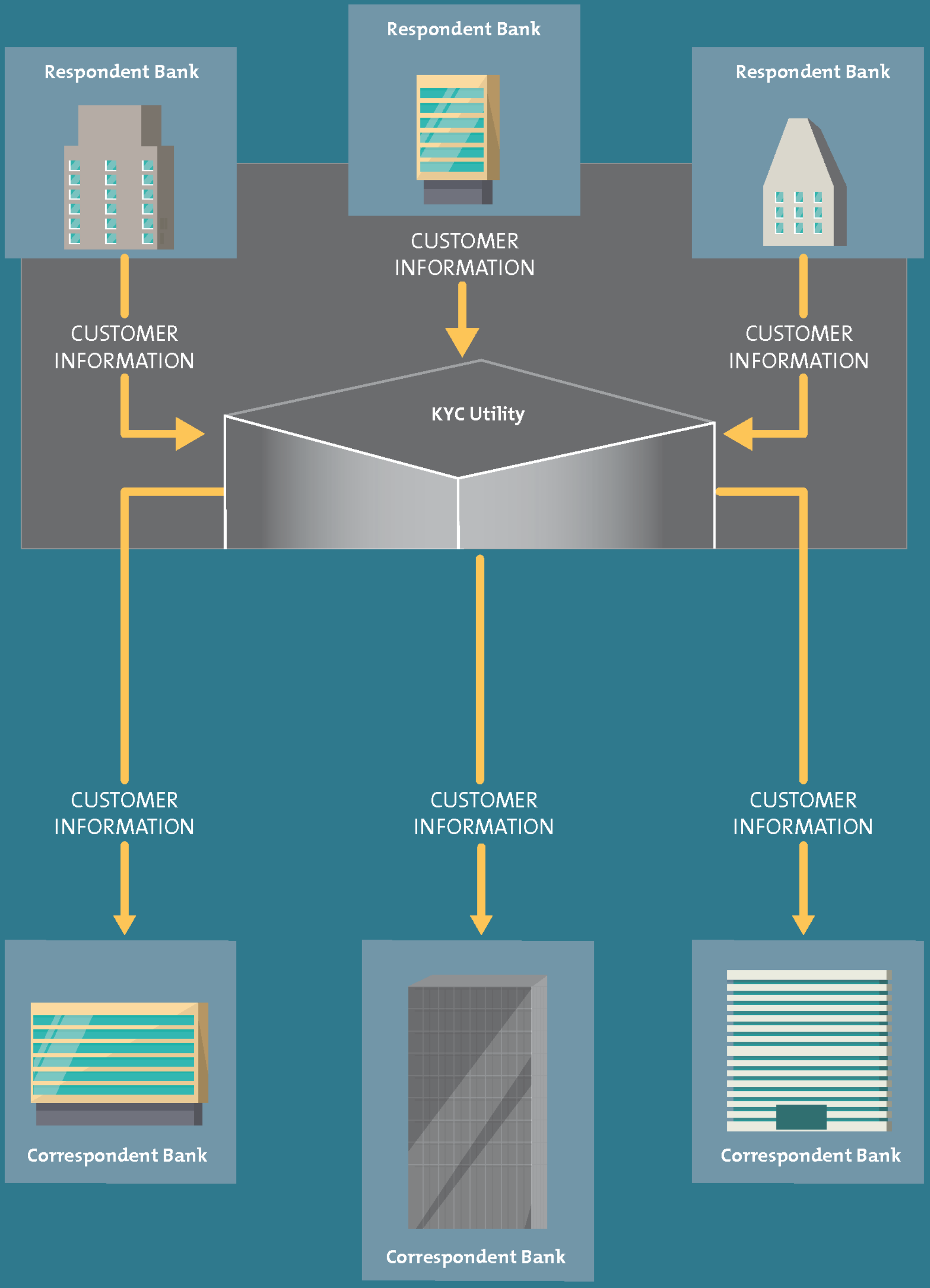
[96] Interview with Bart Claeys, September 8, 2017.

Table 5. Recommendations

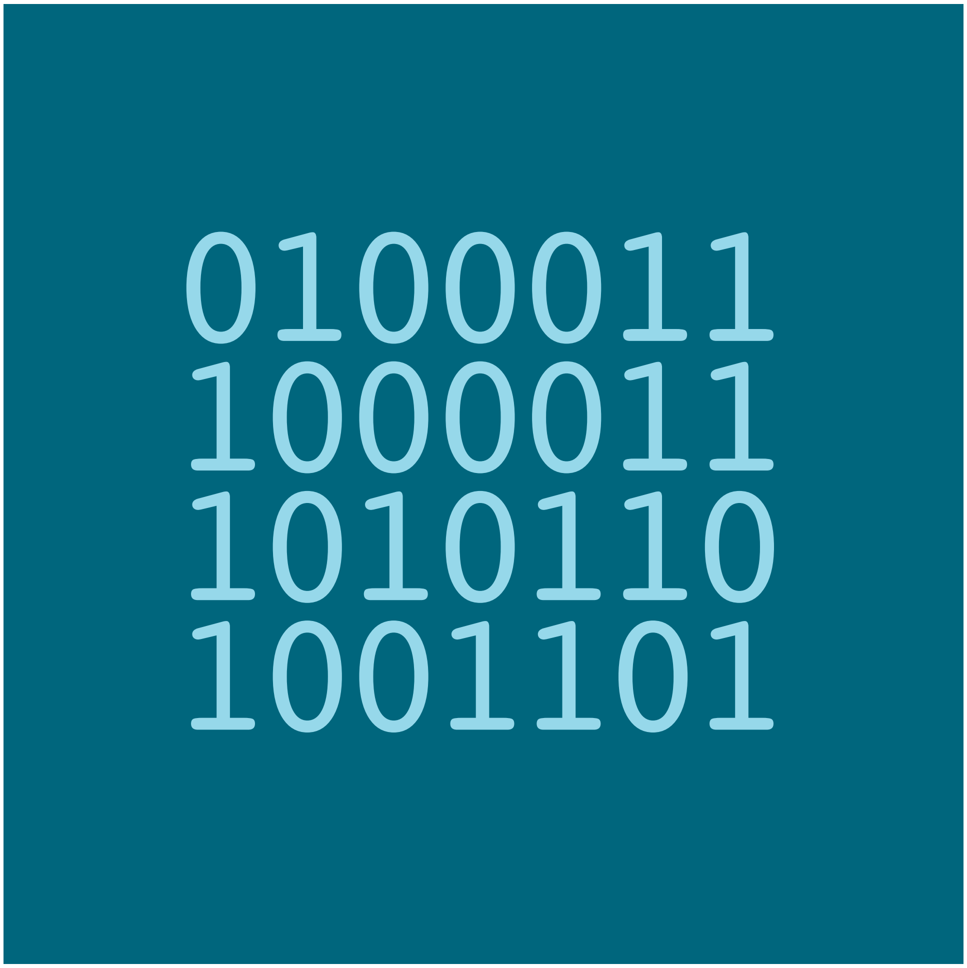
Organizations Involved	Recommendation
National regulators	Give further consideration as to whether and to what degree financial institutions can rely on third parties for customer identification and due diligence, and offer further guidance, if necessary. It is important that banks understand the degree to which they can rely on KYC utilities or other third-party information sharing mechanisms.
National regulators	Provide clarity on who bears (or is allowed to bear) liability if CDD information is incorrect.
National regulators	Consider whether to establish regulatory regimes for regulating and monitoring KYC utilities.
Standard-setting bodies and international organizations	Explore steps necessary to establish KYC utilities that cater to money transfer operators (MTOs) and nonprofit organizations (NPOs). As part of this exploration, work with banks, MTOs, and NPOs to develop standardized due diligence questionnaires similar to the Wolfsberg Group's Correspondent Banking Due Diligence Questionnaire. Also consider whether economic support is needed to make such solutions a reality, or whether they can be market driven.
Standard-setting bodies and international organizations	Continue to engage on developing issues related to KYC utilities.
Standard-setting bodies and international organizations	Explore whether it is possible for third parties to also conduct risk assessments themselves, as opposed to simply providing information for risk assessments.

Know Your Customer Utility

KYC utilities are central repositories for customer due diligence information on financial institutions' direct customers or counterparties



3 BIG DATA



Key points

- *Big data* refers to datasets that are high in volume, high in velocity, and high in variety. Such datasets necessitate different hardware, software, and analytical solutions than traditional datasets.
- Big data is enabled by rapid advances in data capture, transmission, and storage, as well as in computation and analysis.
- The main activities in the big data life cycle include collection, storage, preparation, analysis, and visualization.
- Big data systems can be cheaper, more flexible, and more responsive than traditional relational databases.
- Big data systems can reduce the time compliance staff spend searching for and aggregating information.
- Big data systems expand the scale and scope of data available for KYC and suspicious transaction investigations, enabling more complete and sophisticated analysis than previously possible.
- Big data systems are often linked to advanced analytics processes, including machine learning (discussed more in the next chapter), which may be able to identify patterns and relationships that would otherwise go undetected by human investigators.

What it is

***Big data* is defined as "datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze."**[97] Big data is distinguished by three attributes, referred to as the "three Vs": it has high volume, high velocity, and high variety.[98]

- *High volume:* Big data systems can handle much larger datasets than traditional relational databases.
- *High velocity:* Big data systems can ingest incoming high-volume data in real time.
- *High variety:* Big data systems can gather, process, and store data in many different formats, including semi-structured and unstructured data.

These three characteristics necessitate different hardware, software, and analytical solutions than those used for traditional datasets.[99] Past a certain point, traditional data management systems do not scale efficiently. "If you have more than a million or so rows in a spreadsheet, you probably want to store it in a relational database, such as MySQL," explains Hal Varian, Google's chief economist. "However, if you have several gigabytes of data or several million observations, standard relational databases become unwieldy." For companies that process billions of

[97] Manyika et al., 2011, p. 1.

[98] Laney, 2001. Others have expanded on this definition to suggest numerous other attributes (while keeping to the V theme), including variability, validity, value, and veracity, among others (NIST Big Data Public Working Group, 2015a, p. 7).

[99] NIST Big Data Public Working Group, 2015a, p. 5.

transactions every day, he continues, “analyzing even one day’s worth of data of this size is virtually impossible with conventional databases.”[100]

There are other definitions of big data, but notably, most do not attempt to establish a quantitative standard. Many experts expect that what counts as big data will evolve along with advances in technology.[101] “The ‘big data’ of 15 years ago are most definitely small data by today’s standards,” noted Francis Diebold in 2012. “Moreover, someone reading this in 20 years will surely laugh at my implicit assertion that a 200 GB [gigabyte] dataset is large.”[102]

What counts as big data can vary from one field to the next. In the field of physics, Diebold observed, “200 GB is already small. The Large Hadron Collider experiments that led to discovery of the Higgs boson, for example, produce a petabyte of data (1,015 bytes) per second.”[103]

Big data is broadly enabled by rapid advances in data capture, transmission, and storage, as well as in computation and analysis.

First, the share of digitized information has grown enormously. In 2002, only 25 percent of data was digitized, with the remainder being stored in various analog formats. By 2014, 99.5 percent of stored data was digitized.[104] By 2007, 99.9 percent of transmitted data was digitized.[105] As more and more human and machine activities are intermediated through information technology, and as our ability to capture and store this information grows, the resulting data grows exponentially. “Data is now available faster, has greater coverage and scope, and includes new observations and measurements that were not previously available,” write Einav and Levin, continuing, “Modern datasets also have much less structure, or more complex structure, than the traditional cross-sectional, time-series, or panel data models that we teach in econometrics classes.”[106]

Second, our capacity to transmit and store data has improved rapidly. Between 1986 and 2014, global capacity to store and transmit data increased at a compound annual growth rate of 30 percent.[107] Storage capacity grew from 2.6 exabytes in 1986 to 4.6 zettabytes in 2014, while transmissions capacity grew from 7.5 petabytes in 1986 to 25 exabytes in 2014.[108] (An exabyte is equal to 1 billion gigabytes, or 1,000 petabytes. A zettabyte is equal to 1 trillion gigabytes.)

Finally, our ability to compute and understand this information has grown in tandem with our ability to transmit and store it. Hilbert and Lopez found that for the period from 1986 to 2007, the compound annual growth rate for general-purpose computations was 61 percent; for application-specific computations, the growth rate was 86 percent.[109]

[100] Varian, 2014, p. 4.

[101] Manyika et al., 2011, p. 1.

[102] Diebold, 2012, pp. 1.

[103] Diebold, 2012, pp. 1.

[104] Hilbert, 2015, p. 3.

[105] Hilbert and Lopez, 2012, p. 958.

[106] Einav and Levin, 2013, p. 3.

[107] Hilbert, 2015, p. 3.

[108] Hilbert, 2015, p. 3.

[109] Hilbert and Lopez, 2012, p. 962.

More specifically, big data applications rely on new IT infrastructures, notably the spread of modern data centers, and software advances in areas such as parallel processing. Data centers can house tens of thousands of commodity servers. Parallel processing enables computations to be spread out and coordinated over a number of processors.

Finance is among the most data-intensive industries. McKinsey estimated that in 2009, the securities and investment services sector and, separately, the banking sector had the most stored data, on average, per firm.[110] Combined, the two sectors held the most data of all the sectors examined in the study.[111] Compared with other sectors, financial sectors make especially heavy use of text and numerical data.[112]

How big data works

There are five core functional roles in the big data ecosystem. The National Institute of Standards and Technology identifies them as follows:

- *System orchestrator:* Defines and integrates the required data application activities into an operational vertical system
- *Data provider:* Introduces new data or information feeds into the big data system
- *Big data application provider:* Executes a data life cycle to meet security and privacy requirements as well as system orchestrator–defined requirements
- *Big data framework provider:* Establishes a computing framework in which to execute certain transformation applications while protecting the privacy and integrity of data
- *Data consumer:* Includes end users or other systems that use the results of the big data application provider[113]

At the heart of this ecosystem is the big data application, which ingests, stores, and processes the data, as well as providing a work environment for users to view and analyze the data. The "data life cycle" refers to this process of "transform[ing] raw data into actionable knowledge."[114] Abstracting from particular hardware and software solutions, the big data life cycle may be said to consist of five stages:

1. Collection (also known as acquisition or ingestion)
2. Storage
3. Preparation (also known as transformation or processing)
4. Analysis
5. Reporting and visualization[115]

[110] This is partly due to the size of firms in these two sectors, but it is also a function of the high volume of transactions they process.

[111] Manyika et al., 2011, p. 19.

[112] Manyika et al., 2011, p. 20.

[113] NIST Big Data Public Working Group, 2015d, p. 13.

[114] NIST Big Data Public Working Group, 2015a, p. 8.

[115] NIST Big Data Public Working Group, 2015d, p. 16.

In the first stage, the application's data ingestion engines (also known as *loaders* or *connectors*) pull data (or receive pushes) from the data providers.[116] Data may be sourced from banks' own internal systems or from outside providers. Internal data relevant to AML/CFT compliance work may include client information, transactions metadata, website and app activity, and personal communications (phone calls, e-mails, and so on).[117] External data may be sourced from KYC utilities, public records (such as business registries), traditional or social media, and sanctions and PEPs lists.[118] Initial metadata may be attached at this stage, including sources and subject keys, for later aggregation and indexing.[119] Popular loaders include Apache Sqoop (for batch processing from relational databases), Apache Flume (for streaming log data), and Apache Kafka (also for streaming data).[120]

In the second stage, the data is aggregated in a data pond or a data lake. A *data lake* is a repository for storing raw data. A data lake may be contrasted with a traditional relational database or data warehouse, such as a structured query language (SQL) database. Relational databases require data to be structured or transformed according to preset fields (in effect, columns and rows).[121] The process of storing data and operating on it in a relational database is referred to as "extract, transform, and load" (ETL). In contrast, a data lake stores heterogeneous data in its native format. This can include variably structured data, as well as semi-structured or unstructured data (such as e-mails, documents, phone calls, videos, and so on).[122] Data held in a data lake is structured only "on read"—that is, as requested. The process of storing data and operating on it in a data lake is referred to as "extract, load, and transform" (ELT). One of the most commonly used data lake platforms is Apache Hadoop's Distributed File System.

In the third stage, the data is prepared (or transformed) for analysis. Preparation includes validating the data, cleaning it, and standardizing it; it may also include removing outliers and duplicative entries.[123] Metadata is attached to describe where the data came from (its *lineage*) and any transformations that were performed on it, as well as who can access the data and what operations they can perform on it (its *access control*).[124] Access control may describe rights for human operators as well as software programs.

In the fourth stage, the data is analyzed. Analysis may be performed on a routine basis according to prebuilt reports or in response to ad hoc queries.[125] In contrast to traditional statistical techniques, which rely on sampling, big data analytics "often emphasize the value of computation across the entire dataset, which gives analysts better chances to determine causation, rather than just correlation."[126]

[116] NIST Big Data Public Working Group, 2015d, p. 16.

[117] IIF, 2017a, p. 17.

[118] IIF, 2017a, p. 17.

[119] NIST Big Data Public Working Group, 2015d, p. 16.

[120] Chemitiganti, 2015b.

[121] Mayo, 2016, p. 7.

[122] Mayo, 2016, p. 7.

[123] NIST Big Data Public Working Group, 2015d, p. 16.

[124] Chemitiganti, 2015b.

[125] Oracle, 2013, p. 19.

[126] NIST Big Data Public Working Group, 2015a, p. 16.

These analyses may, in some cases, be performed in real time, as data streams in. Big data analytics may be performed by automated software, such as machine learning algorithms, or with human analysts in the loop, who may employ a variety of techniques for data exploration, hypothesis formulation, and testing.[127] Analytic methodologies include data mining, data fusion, and network analysis, each of which encompasses several discrete techniques, which may be used separately or in combination.[128]

Finally, in the fifth stage, the processed data and analyses are visualized for the data consumer. Visualization can refer to the generation of either text or tabulated reports, as well as to static or interactive graphics.[129] The purpose of visualization is to render the data analytics comprehensible and actionable.[130] Graphical tools may help compliance officers to spot trends or outliers in customer typologies or transactions, as well as to understand how entities are interrelated. Especially useful for AML/CFT compliance are dashboards, heat maps, constellation diagrams, social graphs, and geospatial tools, as are more traditional charts, such as scatterplots and histograms.[131]

Advantages of big data and use cases for AML/CFT compliance

AML compliance staff usually require a wide variety of data to conduct their KYC and suspicious transaction investigations, but in most financial institutions, this information is not available in one place. Rather, the information is spread out across many different organizational silos both inside and outside the organization.[132] This means that the work flow is fragmented and inefficient. Compliance staff cannot conduct their investigations in one place. When investigators receive a suspicious transaction alert in their case management system, they must go out of that system to manually collect the relevant information from various databases and then analyze it. They then write a report, save it as a PDF, and upload it back into the case management system.[133]

In particular, locating, collecting, and aggregating information takes up a great deal of compliance staff's time. According to a survey of AML compliance staff by NextAngles, 34 percent of the time allocated to resolving an AML case is spent on data collection, and another 29 percent is spent on data consolidation and aggregation.[134]

In contrast, data lakes can store all potentially relevant information in one place, reducing investigation times. Because big data systems do not require preset formatting, they can store all types of data, including variably structured data,

[127] NIST Big Data Public Working Group, 2015b, p. 20.

[128] PCAST, 2014, pp. 24-30; Manyika et al., 2011, pp. 27–36.

[129] NIST Big Data Public Working Group, 2015d, p. 17.

[130] NIST Big Data Public Working Group, 2015d, p. 17.

[131] Stabile, 2010, p. 1.

[132] Chemitiganti, 2017.

[133] Chemitiganti and Gillespie, 2016, slide 11.

[134] NextAngles, 2016, p. 16.

semi-structured data, and unstructured data.[135] This means that any information collection and aggregation can be automated, dramatically reducing the time investigators must spend searching for information and streamlining the investigatory workflow.[136]

Another problem is that the data available for analysis in relational databases is often limited, dated, and of uncertain origin. Storage costs for relational databases are expensive. As a result, it is common practice for older data to be archived in secondary or tertiary storage, which limits its accessibility.[137] This means that data available for an investigation or modeling may go back only a few weeks or months.[138] Further, the need to transform data prior to loading it into a relational database is often labor-intensive and leads to long refresh times.[139] Finally, data lineage can be lost in relational databases because raw data is often deleted after the ETL process.[140]

As a result of these data limitations, analysts conducting KYC investigations often limit themselves to a sample of the existing data, rather than taking advantage of all of the information that might be relevant. In particular, analyzing external data (such as news media, social media, and government records) is difficult and is for the most part done manually.

In contrast, big data applications have lower storage costs and faster refresh times. Data lakes can be stored on distributed file systems located in data centers, lowering the cost of storage and improving scalability (both up and down).[141] Lower storage costs also mean that raw data can be retained, improving data lineage issues.[142] Finally, because data can be uploaded in its raw form, big datasets can be updated rapidly, or even in real time.[143]

The ability to use much larger, longer, and more widely sourced datasets vastly expands the analytical possibilities for both CDD and transaction monitoring. By providing efficient access to all available information quickly and easily, big data applications can enable faster and richer forms of analysis, improving responsiveness and accuracy.

Relational databases also have trouble handling complex analyses of large datasets. "Processing scalability" is nonlinear, meaning that as datasets become larger, or as analyses become more complex, batch processing run times become longer and longer; analysts often must choose between richness (running a complex analysis) and reach (using a larger dataset).[144]

Relational databases are also less responsive to new lines of inquiry. Maintaining the ETL layer is "manageable if data sources are all known up front and have unchanging structures, but it can be a major challenge to accommodate new data sources and

[135] Mayo, 2016, p. 8.

[136] Chemitiganti and Gillespie, 2016, slide 14.

[137] Cloudera, 2015, p. 2.

[138] Cloudera, 2015, p. 2.

[139] Mayo, 2016, pp. 6–7.

[140] Mayo, 2016, p. 7.

[141] Mayo, 2016, p. 8.

[142] Mayo, 2016, p. 8.

[143] Mayo, 2016, p. 8.

[144] Mayo, 2016, p. 7.

Table 6. Summary: Advantages and Challenges/Limitations of Big Data

Advantages	Challenges
■ Cheaper, more flexible, and more responsive than traditional database technology ■ Can reduce staff time spent searching for and aggregating information ■ Expands the scale and scope of data available for KYC and suspicious transaction investigations ■ Can be linked to advanced analytic processes	■ May be impeded by regulations that restrict data sharing or mandate data localization

inflexible when dealing with rapid change."[145] Structuring data prior to analysis presupposes the types of questions that will be asked but is less useful for new or exploratory analysis.[146] Moreover, it presupposes the data that will be relevant to a given inquiry. As noted in Booz Allen Hamilton's *Field Guide to Data Science*, "rigid data silos . . . create a filter that lets in a very small amount of data and ignores the rest. These filtered processes give us an artificial view of the world based on the 'surviving data,' rather than one that shows full reality and meaning. . . . Eliminating the need for silos . . . embraces the reality that *diversity is good and complexity is okay*" (emphasis in original).[147] In legacy systems, the need to structure data prior to analysis can slow banks' reaction time to new threats or revelations, such as the Panama Papers leak. In such cases, it can take days or weeks for banks to assess their exposure.[148]

Big data systems more easily allow for customized analysis, as well as novel lines of inquiry. Since big data applications structure data on query, not on loading, they can enable much faster turnaround times on new lines of inquiry—down to seconds or minutes. Such rapid customization facilitates exploratory analysis and allows compliance staff to be much more responsive to new inquiries from management and regulators.

Challenges and limitations

The full exploitation of big data may be impeded by regulations that restrict data sharing.[149] Multinational financial institutions may not be able to aggregate data across all of the jurisdictions they operate in.[150] Impediments to data sharing include bank secrecy laws, privacy and confidentiality laws, and data localization laws. The IIF has argued that such laws make it "complex, if not impossible, for [financial institutions] to obtain a group-wide view of illegal financial activities."[151] In a recent survey of 28 global financial institutions, the IIF reported that three-quarters of respondents felt they were able to share CDD information across their organizations. However,

[145] Mayo, 2016, p. 7.

[146] Mayo, 2016, p. 8.

[147] Blackburn et al., 2015, p. 30.

[148] Interview with David Stewart, director, SAS, August 29, 2017.

[149] IIF, 2017a, p. 17.

[150] IIF, 2017a, p. 12.

[151] IIF, 2017a, p. 12.

Table 7. Recommendations

Organizations Involved	Recommendation
National regulators and international organizations	Determine whether local privacy and data sharing laws pose a challenge to the integration of these datasets and whether these laws can or should be amended without compromising privacy.

respondents "were fairly evenly split" on whether they could share suspicious activity information—and a majority of those who said they could still said they faced limitations on the types of information they could share.[152]

Prospects for adoption

Big data systems will likely become widely adopted in the coming years. The outlook given by Daniel Mayo at Ovum, a technology consultancy, is that Hadoop (a big data application from Apache) "will become a vital data platform in the banking sector for tackling financial crime and compliance. . . . The platform has already matured to the point where it is now being used in live deployment. . . . With most banks actively working with this technology, Ovum expects significant adoption [over] the next two to three years."[153]

Assessments by regulators, policymakers, standard-setting bodies, and trade associations

The International Monetary Fund (IMF) has given brief favorable mention to big data. In a recent paper on fintech (financial technology), the IMF devoted a paragraph to regtech, noting big data's potential to identify suspicious transactions.[154] In addition, in a recent speech, IMF Managing Director Christine Lagarde gave a positive nod to regtech, calling it "a powerful tool" against terrorist financing: "We can use fintech to identify terrorist financial flows, including in the case of very small transactions. Machine learning and artificial intelligence tools can help identify patterns of activity that would otherwise be difficult to detect" (emphasis in original).[155]

Other international organizations are beginning to consider how to address big data for AML/CFT compliance. The FSB's Correspondent Banking Coordination Group has held an initial discussion on big data and advanced analytics, and is considering whether to pursue the issue further.[156]

US policymakers and regulators have not commented publicly on big data's applicability to AML/CFT compliance. However, more generally, the US Securities and Exchange Commission, among financial regulatory agencies, has been at the forefront of publicly considering how to use big data to tackle fraud and other financial crimes.[157]

[152] IIF, 2017b, p. 8.

[153] Mayo, 2016, p. 1.

[154] He et al., 2017, p. 18.

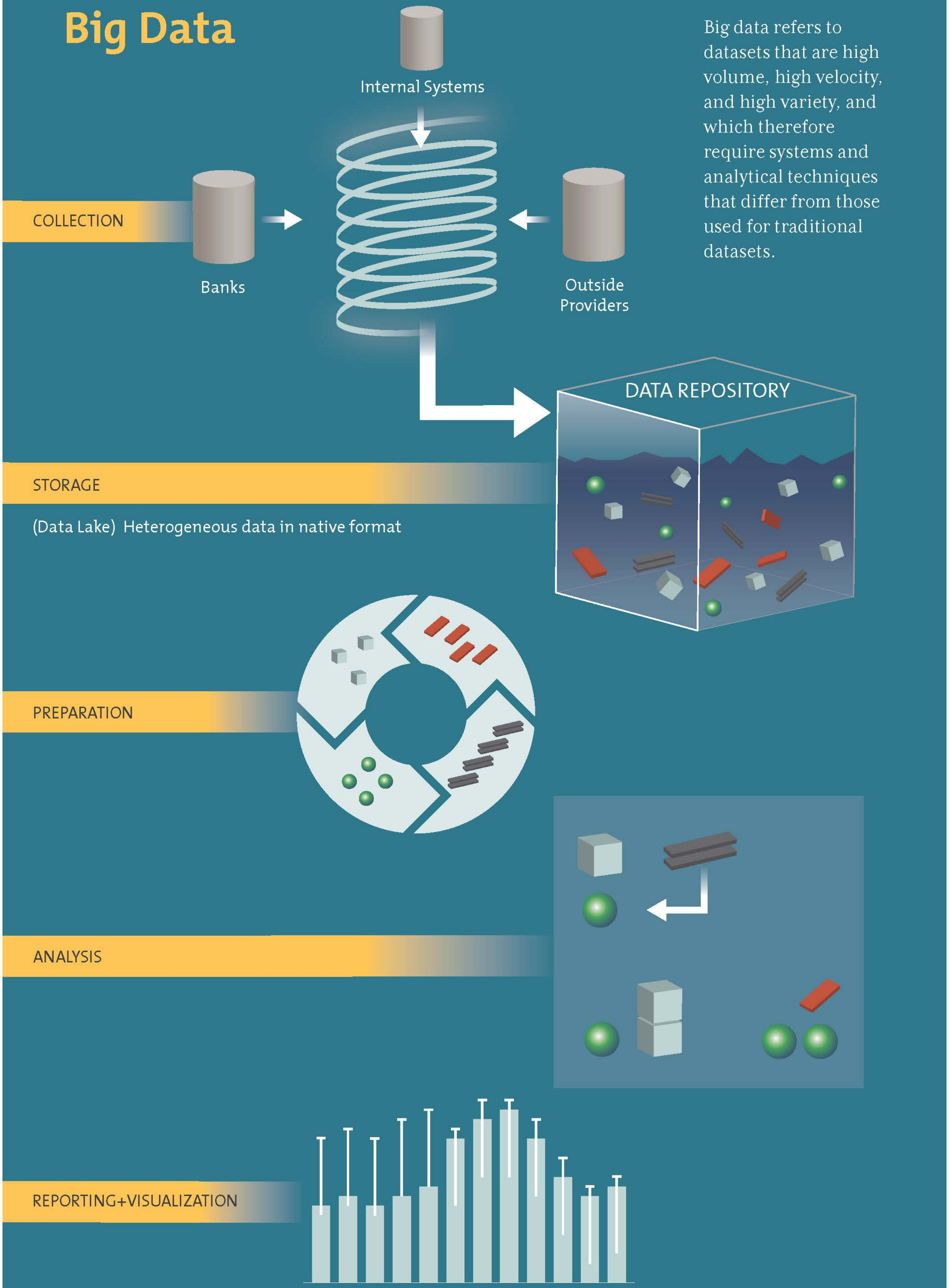
[155] Lagarde, 2017.

[156] FSB, 2017b, p. 21.

[157] Stein, 2016; Bauguess, 2017.

Big Data

Big data refers to datasets that are high volume, high velocity, and high variety, and which therefore require systems and analytical techniques that differ from those used for traditional datasets.



Internal Systems

COLLECTION

Banks

Outside Providers

DATA REPOSITORY

STORAGE

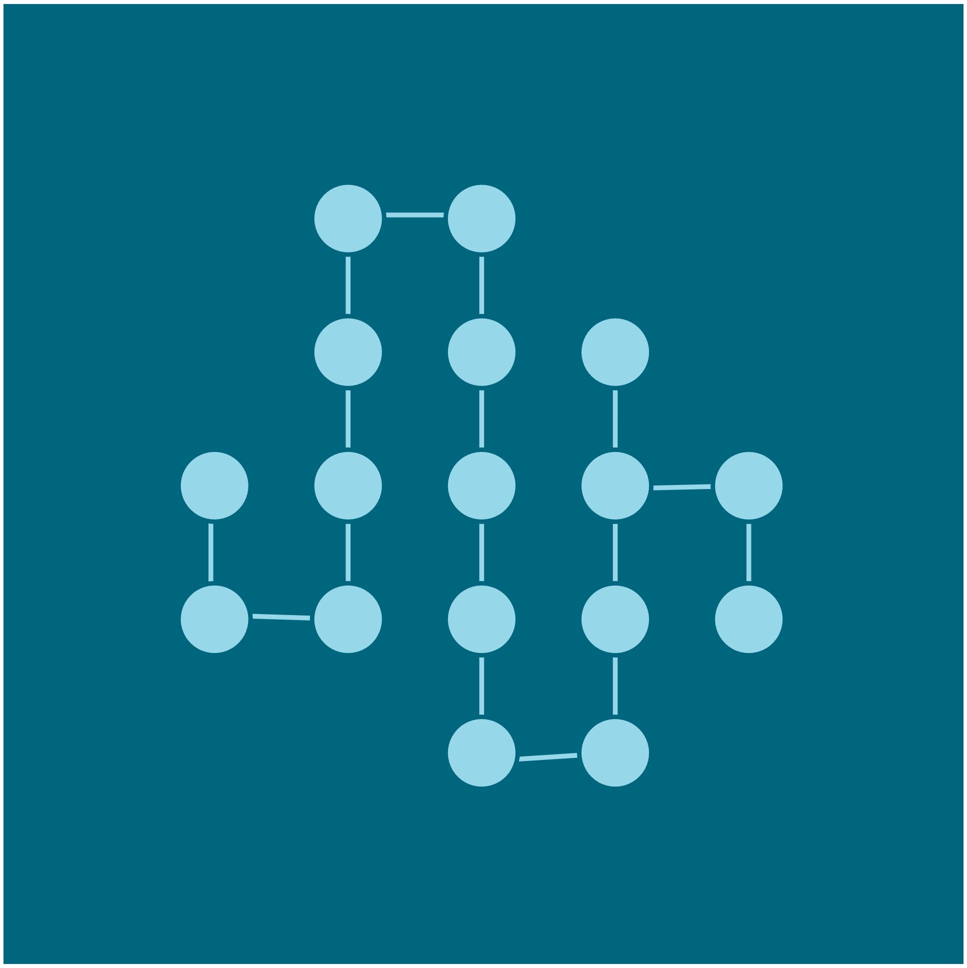
(Data Lake) Heterogeneous data in native format

PREPARATION

ANALYSIS

REPORTING+VISUALIZATION

4 ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING



Key points

- Machine learning is a type of artificial intelligence (AI), which itself is a branch of computer science.
- Machine learning enables computer programs to improve their performance at a given task through repeated iterations.
- Supervised machine learning uses labeled datasets to develop models that can accurately predict a predefined output.
- Unsupervised machine learning explores unlabeled data, searching for patterns and relationships.
- Machine learning works best with large datasets and therefore has benefited from the emergence of big data and advances in computation.
- Machine learning can be used to help financial institutions reduce false negatives in suspicious activity alerts by developing more accurate models for detecting illicit finance.
- Anomaly detection, a type of unsupervised machine learning, can be used to detect illicit finance techniques that were previously unknown to banks and government authorities.

What machine learning is

Machine learning is a type of AI (itself a branch of computer science) whereby a computer program is able to improve its performance at a given task through repeated iterations. Arthur Samuels, a pioneering computer scientist who worked at Bell Labs, IBM, and Stanford University, first explained in 1959 that "programming computers to learn from experience should eventually eliminate the need for much . . . detailed programming effort."[158] Tom Mitchell, a computer scientist at Carnegie Mellon, put forth a more formal definition in 1997: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E."[159]

Though first introduced in the late 1950s, machine learning was not widely applied until the 1990s, following disappointments with alternative approaches to AI, such as microworlds and expert systems.[160] Previous techniques "used logical rules to model intelligence," explains Andreessen Horowitz, a venture capital firm, in its guide to artificial intelligence.[161] "Building AI meant representing the world in a set of data structures (such as trees or lists or sets) and then using rules (such as *and*, *or*, *if-then-else*, and so on) to reason about that knowledge."[162] These approaches had some successes in very constrained environments but lacked extensibility. Outside of the constrained environments for which they were built, these AI programs would quickly break down, becoming nonsensical.[163] In contrast, machine learning allows

[158] Samuel, 1959, p. 535.

[159] Mitchell, 1997, p. 2.

[160] Chen, 2016.

[161] Andreesen Horowitz, 2017.

[162] Andreessen Horowitz, 2017.

[163] Chen, 2016.

computers to “derive their own ‘rules’ using lots and lots of data.”[164] Machine learning demonstrated that it could be efficiently applied to a wide range of problems.

Machine learning has been accelerated by three advances: greater availability of data, more powerful computers, and better algorithms. Because machine learning works best on large datasets, it is “closely associated with” big data.[165] More powerful computers mean that more computing cycles can be run on the data in less time, accelerating the learning process.[166] The successes of machine learning have also led to a surge of public- and private-sector investment, which has further accelerated progress.[167]

Of particular promise is deep learning, an advanced form of machine learning. Deep learning uses networks of machine learning algorithms. Deep learning can be used to analyze unstructured data, including non-numerical data, such as photos and video.[168] However, deep learning requires even larger datasets than traditional machine learning.

Today, applied AI is “narrow,” as opposed to “general.” AI models must be trained to carry out specific tasks, and each task needs to be trained separately. General AI, which would allow computers to learn and make decisions across multiple domains, remains out of reach.[169]

How machine learning works

There are three broad types of machine learning—supervised learning, unsupervised learning, and reinforcement learning. These are distinguished by the type of feedback mechanism. With supervised learning, the machine learning algorithm builds a statistical model to predict or estimate a predefined output based on the available inputs. With unsupervised learning, the machine learning algorithm explores a dataset for patterns and relationships, without a predefined output. Reinforcement learning falls between the two, with the algorithm receiving general feedback on its performance, but without a specific predefined output to aim for.

Supervised learning relies on training a model with labeled data. The process begins with creating a dataset and labeling the data points according to the predefined output. For example, if we were to develop a model that could accurately recognize suspicious transactions, we would first construct a dataset of past transactions and label these as “suspicious” or “not suspicious,” perhaps according to whether the transaction resulted in the filing of a suspicious activity report (SAR) or not. This dataset would then constitute the training set, which the program would use to develop a model. The program would then test the model against the labeled data and compare its own answers with the true answers. The program would then iteratively revise the model to close the gap between its own answers and the true answers.[170] The program would then test its model on out-of-sample

[164] Andreessen Horowitz, 2017.

[165] Van Liebergen, 2017, p. 61.

[166] Chen, 2016.

[167] Chen, 2016.

[168] van Liebergen, 2017, p. 63; IIF, 2017a, p. 19.

[169] Petrasic et al., 2017, slide 7.

[170] Andreessen Horowitz, 2017.

data, with its judgment being assessed by programmers. Supervised learning may be used to predict continuous outputs (as in the case of regression analysis—sometimes referred to as *numerical prediction* in the machine learning field) or a discrete output (as in classification analysis).

Unsupervised learning is used when the answer is not known or when the datasets are unlabeled. Unsupervised learning is useful for organizing data into groups based on common characteristics. This is known as clustering and is similar to classification but used for unlabeled data. Unsupervised learning may also be used to develop customer typologies based on common behaviors or transaction patterns. Unsupervised learning may also use these customer typologies to flag outliers, a process known as anomaly detection.

Of special relevance to finance—and, in particular, to AML compliance in correspondent banking—is feature engineering. In recent years, much of the excitement about machine learning has centered around machine perception capabilities, which are used to interpret incoming sensor data.[171] Machine perception underlies speech recognition technologies and is also a critical component of self-driving cars. However, in finance, many of the issues that machine learning is applied to involve the analysis of data that is sparse (i.e., it has many missing values) and highly dimensional.[172] That is, the information a bank has on two different customers may be very different—even the types of data it has may be different.[173] In a large dataset with 10,000 columns, for example, many rows may be blank. To address this issue, machine learning may use feature engineering to fill in the missing data.[174] This is especially useful when engaging in transaction monitoring, in which the information on payment originators and beneficiaries is often incomplete or missing, especially when coming from developing countries.[175]

Problems machine learning addresses

Machine learning may help financial institutions to reduce false negatives and false positives in suspicious activity alerts.

False positives are a major issue—if not the major issue—in AML/CFT compliance. Most investigations do not result in an SAR, meaning that a large portion (perhaps the majority) of time spent by banks' investigations teams is not well spent.[176]

The traditional, rules-based approach to transaction monitoring is time-consuming, vulnerable to bias, and simplistic. In the traditional, rules-based approach, compliance officers collect KYC information and review that and transactions against rules that describe what constitutes suspicious behavior.[177] These rules delineate the patterns that are thought to indicate illicit transactions.[178] Typically, money-laundering patterns are "hand coded" (manually classified) by subject-matter experts and

[171] Interview with Gurjeet Singh, executive chairman and co-founder, Ayasdi, August 18, 2017.

[172] Interview with Gurjeet Singh, August 18, 2017.

[173] Interview with Gurjeet Singh, August 18, 2017.

[174] Interview with Gurjeet Singh, August 18, 2017.

[175] Interview with Gurjeet Singh, August 18, 2017.

[176] Pasquali, 2017.

[177] Petrasic et al., 2017, slide 8.

[178] Readling, 2016.

are divided by geography and business type or customer segment. Then, when one of these rules is triggered, the flagged transaction must be investigated.

The hand-coded patterns focus on individual transactions or simple transaction patterns.[179] They are unable to spot complex transaction patterns or to take a systemwide view of transaction behaviors.[180]

Moreover, these processes are carried out separately for customers and for transactions.[181] This means that the analysis does not take a holistic view of the customer. Moreover, because these categories are fixed and relatively coarse, they capture a large number of legitimate transactions, resulting in a high number of false positives.[182]

Financial institutions often incorporate new regulatory requirements by "layering" them on top of old requirements, in order to ensure that nothing is missed. This practice is inefficient, however, as it can lead to redundancy in compliance. AI may help compliance officers to understand where requirements overlap, thereby ensuring full compliance without redundancy.[183]

False negatives may also be a significant—though less appreciated—issue. Money laundering techniques are not static. They evolve over time in response to AML/CFT techniques. AML/CFT controls need to be dynamic as well.

Financial institutions do not always know how to identify money laundering or terrorist financing. They rarely receive government feedback on the SARs they file. Moreover, money laundering techniques evolve over time as criminals and terrorists adapt their behavior and methods to avoid detection, and financial institutions are sometimes slow to adapt their surveillance and detection practices. Because it operates with a lag, a rules-based approach will fail to detect these evolving techniques in a timely fashion.

Advantages of machine learning and use cases for AML/CFT compliance

Machine learning programs do not rely on codified rules and can include a wide array of variables in their analysis—as many as available data allow. According to Petrasic, Saul, and Bornfreund, such variables may include:

- Where a customer opens an account relative to [his or her] home address
- What time of day an account was opened
- Duration between transactions
- Patterns among merchants where a customer makes transactions
- Relationships between other customers of those same merchants
- Whether a customer uses a mobile telephone
- What communication channel a customer uses to contact the bank
- Changes in the customer's social media presence[184]

[179] van Liebergen, 2017, p. 65.

[180] IIF, 2017a, p. 18.

[181] Ayasdi, 2017, p. 2.

[182] Ayasdi, 2017, p. 2.

[183] Sparks, 2017.

[184] Petrasic, Saul, and Bornfreund, 2017, p. 5.

An AI system may be able to observe patterns or relationships among the above data that human analysts would have difficulty seeing.[185] It is enormously difficult to consistently spot sophisticated money laundering and terrorist finance activities, since the perpetrators are focused on blending in. "You're not looking for a needle in a haystack," explains Simon Moss, then a managing director at Grant Thornton. "You're looking for a needle in a stack of needles."[186]

Machine learning algorithms that search for anomalous behavior may be used to detect new money laundering techniques. Traditional rules-based transaction monitoring systems are built around known behaviors, such as structuring. Sophisticated money launderers and financers of terrorism are constantly experimenting with new ways of evading detection, making the threat environment a dynamic one. New money laundering techniques may include digital charities, digital gaming and retailing, special investment vehicles, and high-end real estate transactions.[187] Machine learning algorithms can be trained to understand what constitutes a normal transaction for a particular type of originator or beneficiary, and then to flag anomalies. David McLaughlin, CEO of QuantaVerse, describes a situation in which one of his company's programs flagged a transaction between a computer manufacturer and a casino. While casinos are major purchasers of IT equipment, the machine learning program saw that the money was flowing from the manufacturer to the casino, not the other way around. Because the algorithm had never seen this before, it flagged the transaction for further investigation, whereupon suspicious activities were discovered.[188]

Cluster analysis, a type of unsupervised learning, may be used to better segment clients and counterparties. Accurate segmentation is necessary to determine what constitutes normal behavior for a given customer and what may be suspicious. However, traditional customer taxonomies are static, meaning they may miss evolving traits and behaviors, resulting in false negatives. They are also overly broad, resulting in too many false positives.[189] Cluster analysis, in contrast, can allow for segmentation along many more dimensions, resulting in much finer segmentation and a reduction in false positives. Ayasdi, an AI firm, was able to reduce KYCC investigations at one major correspondent bank by 25 percent while simultaneously discovering previously undetected risks.[190] The firm's AI technology did so by first identifying more than a thousand features it could use to distinguish payment originators and beneficiaries and then selecting 120 of these for customer segmentation. The bank's legacy transaction monitoring system only used 10 features.[191]

Machine learning can improve compliance teams' analytical abilities.[192] It is "highly efficient for exploring high-volume or high-dimensional data."[193] It can be used to organize and analyze large datasets—especially variably structured datas-

[185] Petrasic, Saul, and Bornfreund, 2017, p. 5.

[186] Interview with Simon Moss, August 25, 2017.

[187] Interview with Simon Moss, August 25, 2017.

[188] Interview with David McLaughlin, founder and CEO, QuantaVerse, August 25, 2017.

[189] Ramachandran, 2016.

[190] Ayasdi, 2017, p. 4.

[191] Ayasdi, 2017, p. 5.

[192] IIF, 2017a, p. 18.

[193] IIF, 2016, p. 12.

Box 4. Machine Learning for Government Agencies

The use of machine learning to analyze patterns of money laundering and terrorist financing need not be limited to financial institutions. Government agencies would benefit from utilizing these systems, too.

The government could use the systems to sift through SARs filed by banks. After the passage of the USA PATRIOT Act, the rate of growth in SAR filings increased by 55 percent.[1] (This may be due to [1] the USA PATRIOT Act's increasing the number of entities required to file SARs or [2] concerns about reputation risk and fines, leading banks to err on the side of filing.[b])

- FinCEN could use AI to analyze its database of SAR reports against its records of AML/CFT investigations.[3] This could lead to faster resolution of AML/CFT investigations and improve Treasury's understanding of the evolving methodologies used by terrorists and criminal organizations.
- If Treasury could share with banks which SARs led to investigations or even prosecutions, this would enable AI systems to improve further.

a. Klein and Readling, 2015.

b. Klein and Readling, 2015.

c. Readling, 2016.

ets collected by big data systems—relevant to KYC investigations and transaction monitoring.

In particular, machine learning is useful for identifying nonlinear patterns in data. Nonparametric machine learning algorithms can select the most appropriate type of function (e.g., linear, exponential, power, etc.) for a given dataset.[194]

In transaction monitoring, fuzzy matching may be used for entity resolution of payment originators and beneficiaries.[195] Often, payment messages' information on originators and beneficiaries is incomplete. Fuzzy matching uses probabilistic modeling to determine the likelihood that a given value matches another. It can be used to augment the screening of transactions for sanctions and PEPs.[196]

Machine learning can improve the accuracy of its decisions/predictions by updating its model as new data comes in.[197] This also makes it better at detecting changing behaviors.

Machine learning may be used to determine how often to conduct due diligence checks on current customers.[198] The current practice is to do periodic due diligence checks at set intervals based on the perceived riskiness of the client, or when some information about the client changes. Such checks are time-consuming and expensive. Machine learning could be used to determine the most appropriate time to conduct due diligence on a customer.

[194] van Liebergen, 2017, p. 67.

[195] Interview with David Stewart, August 29, 2017.

[196] Cognizant, 2014, p. 6.

[197] IIF, 2016, p. 18.

[198] Interview with Gurjeet Singh, August 18, 2017.

Box 5. Industry Developments

In 2016, IBM acquired Promontory Financial Group, a Washington, DC–based consulting firm, with the goal of merging the group's expertise in financial regulatory compliance with the capabilities of IBM's Watson Financial Services. Promontory is in the process of teaching Watson about financial regulations.[a]

Watson is designed to be evidence based, to allay concerns that it is a black box. Instead of just providing an answer, it provides its reasoning and evidence as well.[b]

[a] Sparks, 2017.
b Sparks, 2017.

AI systems could also help banks to identify attempts to circumvent controls imposed by the US Department of the Treasury's Office of Foreign Assets Control. Current systems largely rely on individuals to identify themselves truthfully.[199]

AI can also be used to power automated compliance programs, which can allow compliance teams to handle much larger investigative volumes. Automated compliance programs can handle the more rote aspects of investigations, freeing up human investigators. Because investigations are time-consuming and costly, only a fraction of alerts are currently investigated.[200] If AI can be used to reduce the time it takes to conduct an investigation from hours to seconds, then many more alerts can be investigated, potentially uncovering instances of money laundering that previously would have been passed over.[201]

Challenges and limitations

Currently, machine learning models need to be designed, built, and trained by humans; they cannot yet do these tasks themselves.[202] Machine learning models cannot transfer what they have learned to new tasks, even if similarities or overlaps exist. (For instance, Google's AlphaGo machine cannot play chess.)[203]

Designing and training machine learning models is time-consuming and requires large amounts of data. Mastering tasks that are simple for humans can require millions of data points and a lot of manual training for machines.[204]

Financial institutions do not always have sufficient high-quality data to train machine learning algorithms.[205] Supervised machine learning algorithms need to be trained with clearly labeled historical data (e.g., with a dependent variable of "money laundering" versus "not money laundering"). However, financial institutions rarely receive government feedback on their SARs, so they do not know which

[199] Petrasic, Saul, and Bornfreund, 2017, p. 5.

[200] Interview with Simon Moss, August 25, 2017.

[201] Interview with Simon Moss, August 25, 2017.

[202] Windsor, 2017, p. 8.

[203] Windsor, 2017, pp. 7–8.

[204] Windsor, 2017, p. 7.

[205] van Liebergen, 2017, p. 65.

Table 8. Summary: Advantages and Challenges/Limitations of Machine Learning

Advantages	Challenges
■ Can reduce false positives in transaction monitoring, perhaps dramatically ■ Can identify previously undetected illicit transactions ■ Can reduce costs by automating certain transaction monitoring tasks currently performed by analysts	■ Better at prediction than establishing causation; may present a trade-off between accuracy and explainability ■ Requires large amounts of high-quality data to train models

ones were true positives and which ones were false positives.[206] As a result, financial institutions may have to rely on historical data with "second-best" dependent variables (e.g., "SAR" versus "not SAR") to train their algorithms.[207] The lack of clearly-labeled data can be addressed to some extent with unsupervised learning, which does not require labeled data. Instead, the AI-based system identifies patterns on its own.

Even if financial institutions do have adequate data, they need to manage their data very carefully. Financial institutions need to understand what data they are permitted to collect, who has rights to it; how it must be anonymized; where it can be stored; and how quickly it must be corrected or deleted.[208]

Financial institutions need to be aware of biases that can affect their machine learning programs. These include input bias, programming bias, and training bias.[209] When AI systems are used to screen domestic customers, such biases could result in consumer protection violations.[210] Even in areas where consumer protection laws do not apply, biases still matter, as they can negatively affect the detection capabilities of the system, raising either false positives or false negatives.[211]

Machine learning can be viewed as a black box, which is unacceptable to regulators. Not only can it be difficult to understand why particular decisions were made, but it may also be difficult to spot errors or biases that have crept into the model. Several compliance officers have said they decided not to acquire AI monitoring tools because they were concerned regulators would view them as a black box, which would limit auditability.[212] In contrast, banks have successfully incorporated AI-based monitoring systems into their antifraud functions, but they have been able to do this in part because regulatory oversight is not nearly so stringent—PayPal, for instance, was able to reduce false positives by half.[213] The black box issue is being addressed by the development of AI programs that explain the reasoning behind their recommendations.

[206] IIF, 2017a, p. 19.

[207] van Liebergen, 2017, p. 66.

[208] Petrasic et al., 2017, slide 17.

[209] Petrasic et al., 2017, slide 21.

[210] Petrasic et al., 2017, slide 21.

[211] Interview with Kevin Petrasic and Benjamin Saul, August 7, 2017.

[212] Bethencourt, 2017.

[213] Bethencourt, 2017.

Table 9. Recommendations

Organizations Involved	Recommendation
National regulators	Share feedback on SAR submissions.
National regulators	Allow financial institutions to share data so as to expand the pool that machine learning programs can learn from.
National regulators	Consider a regulatory sandbox to allow financial institutions to experiment with machine learning solutions.

KYC regulations are rigidly prescriptive and may not allow for much process innovation. KYC may be considered a rigid compliance regime, in which case it may be better to continue to perform it manually, even if AI is more efficient.[214] An AI system that was designed to follow the rules-based approach would replicate its inefficiencies—generating the same false positives that banks already are drowning in.[215]

Prospects for adoption

Large financial institutions will be the first to adopt machine learning for AML/CFT compliance. They are likely to derive the greatest return on their investment, owing to their large volumes of transactions as well as the complexity that arises from their geographic reach and multiple lines of business. They are also best placed to incorporate AI systems, being the most technologically advanced and having the largest amounts of data at their disposal to train AI models.[216]

Mid-size banks may have much to gain from the dispersion of machine learning technologies. Although they will not necessarily be the first adopters, mid-size banks have as much to gain as the large institutions, if not more. Mid-size banks cannot afford the large compliance staffs that large financial institutions have, and so they struggle under mounting compliance expectations. Machine learning may help to place mid-size banks on more competitive footing.[217]

Banks have begun running regtech systems in parallel to traditional KYC and AML systems, in the hopes that doing so will demonstrate the greater efficacy of these systems and persuade regulators to allow them.[218]

Due to criminal, legal, and reputational risks, AI systems have to be extremely robust before being made operational. This precludes iterative experimentation and means that adoption may be slow.[219] Programming errors can lead to disastrous consequences for banks in the area of AML/CFT compliance.

[214] Petrasic et al., 2017, slide 20.

[215] Bethencourt, 2017.

[216] Ray and Katkov, 2016, p. 17.

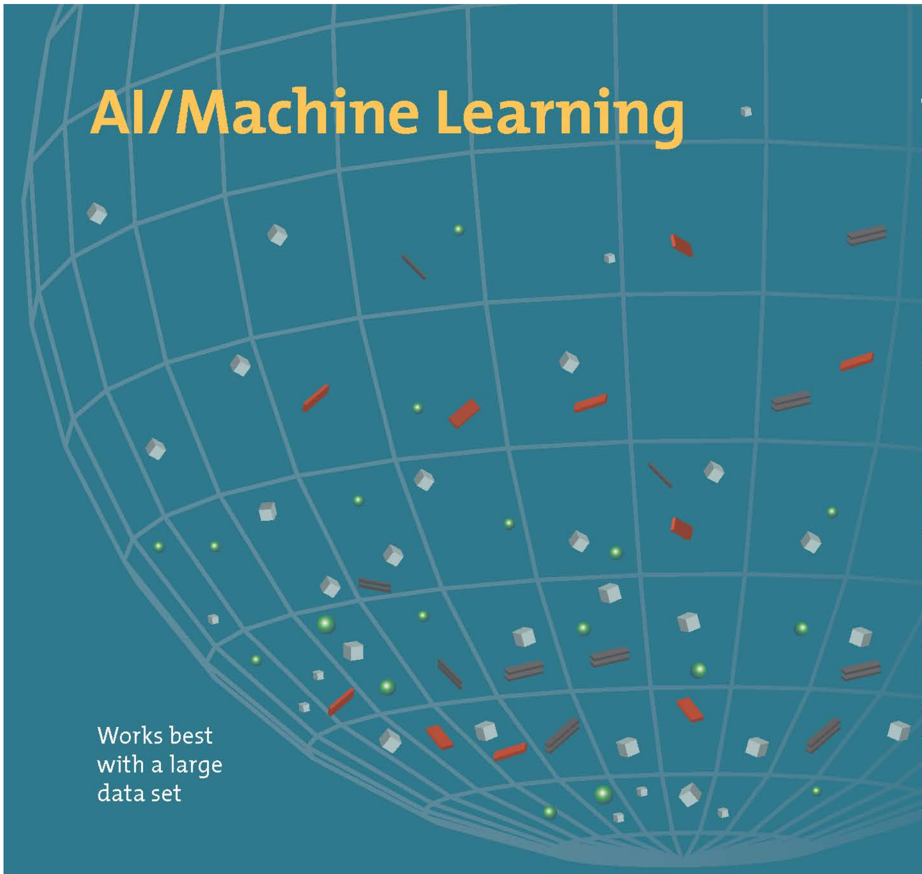
[217] Interview with David Stewart, August 29, 2017.

[218] Petrasic, Saul, and Bornfreund, 2017, p. 6.

[219] Windsor, 2017, p. 8.

AI/Machine Learning

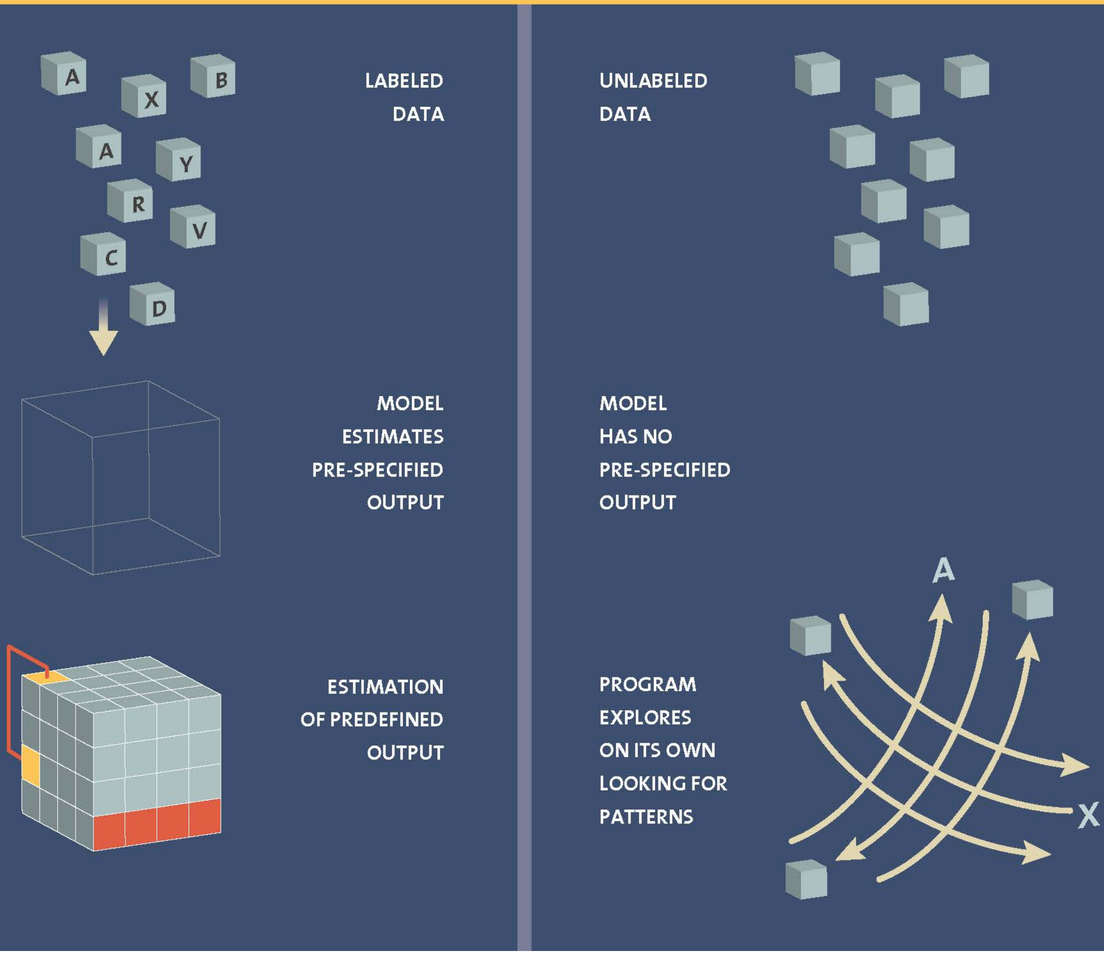
There are three broad types of machine learning—supervised, unsupervised and reinforcement learning. With supervised learning, the machine learning program analyzes a dataset to build a model that best predicts a predefined output. By contrast, with unsupervised learning, the machine learning program is not given a predefined output—rather, it explores the data on its own, looking for patterns and relationships in the dataset. Reinforcement learning falls between the two, with the algorithm receiving general feedback on its performance, but without a specific predefined output to aim for.



Works best with a large data set

SUPERVISED MACHINE LEARNING

UNSUPERVISED MACHINE LEARNING



LABELED DATA

MODEL ESTIMATES PRE-SPECIFIED OUTPUT

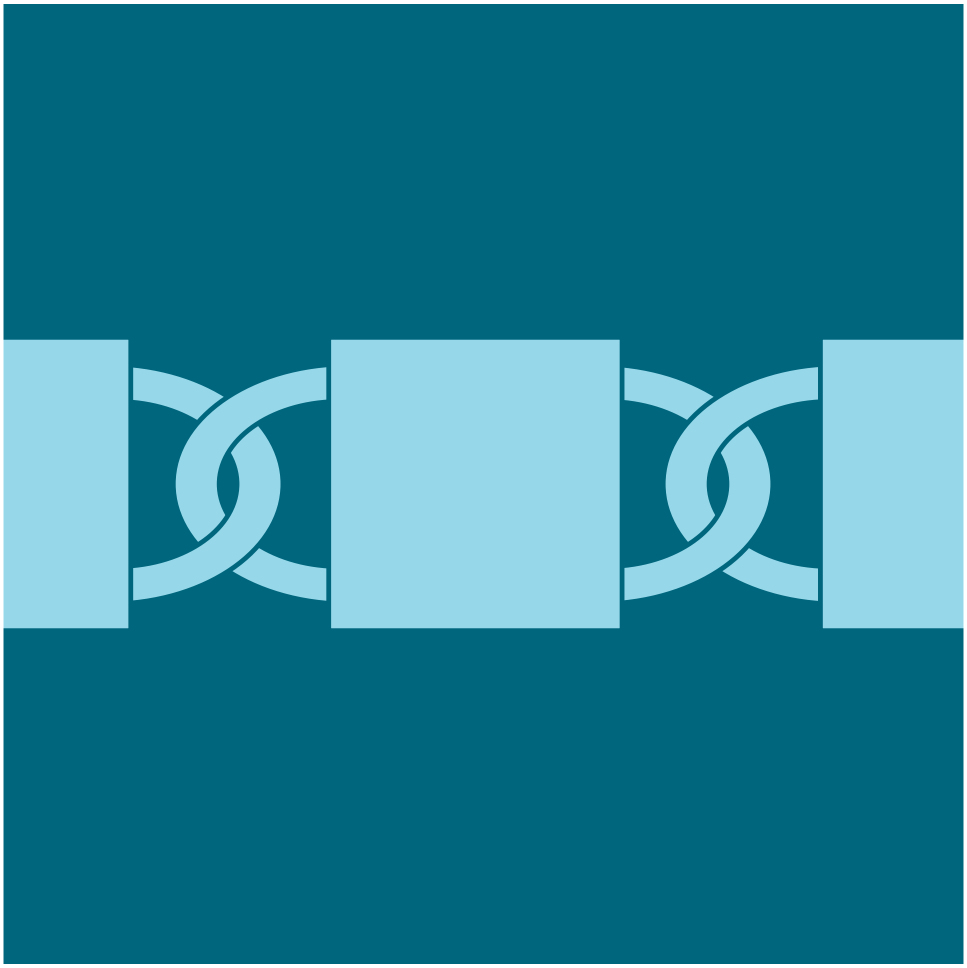
ESTIMATION OF PREDEFINED OUTPUT

UNLABELED DATA

MODEL HAS NO PRE-SPECIFIED OUTPUT

PROGRAM EXPLORES ON ITS OWN LOOKING FOR PATTERNS

5 DISTRIBUTED LEDGER TECHNOLOGY AND BLOCKCHAIN



Key points

- Distributed ledger technology (DLT) is used to securely manage data on a peer-to-peer (P2P) network of computers.
- DLT's defining features include P2P networking, synchronized data sharing, consensus-based governance, and cryptography.
- Blockchain is a type of DLT in which data modifications are recorded in time-stamped "blocks," each of which is connected to the previous block, forming a chain.
- DLT arrangements may be open or closed. They may also be permissionless or permissioned. Financial institutions are mainly interested in closed, permissioned arrangements.
- DLT may benefit regulatory compliance by enabling secure record keeping and instantaneous information sharing.
- DLT could be useful for storing and sharing KYC information, particularly on natural persons.

What DLT is and how it works

Blockchain is a type of DLT. DLT refers to synchronized digital databases that are replicated across a network of computers. DLT makes it possible to store, share, and transfer information securely without a central administrator.[220]

Although DLT is commonly associated with virtual currencies, it has many other applications. A ledger is simply a type of database. It may be used to record the ownership of and transactions with digital assets (monetary or not), but it can also be used to store, share, or exchange any other type of data.

DLT arrangements run on P2P networks. Most databases run on centralized networks, also called client-server networks, in which a central hub (the *server*) "acts as a single source of valid information and control" for the other nodes in the network (the *clients*).[221] In contrast, P2P networks operate without a central hub—the nodes share responsibility for managing the database. On the most decentralized P2P networks, the nodes are equal and undifferentiated—they can function as clients and servers simultaneously.[222]

P2P networks may be managed by a single entity, or they may comprise many distinct entities.[223] These entities may include individuals, financial institutions, and government regulators, among others.[224] As with any other Internet-enabled technology, DLT networks can span multiple jurisdictions.[225]

[220] CPMI, 2017, p. 2.

[221] Mills et al., 2016, p. 10.

[222] Mills et al., 2016, p. 10.

[223] CPMI, 2017, p. 7.

[224] GAO, 2017, p. 40.

[225] CPMI, 2017, p. 3.

DLT arrangements may be *open* or *closed*. Open ledgers are accessible to everyone, whereas closed ledgers restrict their membership to entities that are preapproved or else meet certain criteria.[226]

DLT arrangements may be *permissioned* or *permissionless*. This distinction refers to whether or not the arrangement differentiates nodes by function. Permissionless arrangements allow nodes to assume any and all functions on the ledger. In contrast, permissioned arrangements assign nodes to specific roles. For example, some nodes might be allowed to issue new digital assets or to propose updates (state changes) to the ledger, while others are entrusted with validating those issuances and updates. Yet other nodes may serve as administrators, and still others may serve as auditors and be restricted to read-only status.[227]

For reasons of privacy, security, and control, financial institutions exploring DLT applications are mainly interested in closed, permissioned arrangements.[228] Closed, permissioned networks may also be more efficient and have higher throughput, because they can rely on validation protocols that are less computationally intensive than those typically used by open, permissionless systems.

DLT makes extensive use of cryptography. Cryptography is used for identification, authentication, and permissioning, as well as for data encryption, validation, and time stamping.[229]

The entire ledger need not be transparent to all participants. Even if every node retains a complete copy of the ledger, it is possible to encrypt the data so that each node's users can view only the information that is relevant to them.[230]

Distributed ledgers operate by consensus. A DLT arrangement's consensus mechanism defines the protocol by which updates to the ledger are proposed, validated, and accepted.[231] Individual nodes may propose updates to the ledger, but these updates will be accepted only if they are first validated and agreed to by a specified quorum of nodes on the network. The basic procedure for updating the ledger runs as follows:

1. *Proposal:* A node proposes an update to the ledger.
2. *Validation:* One or more validators authenticate the identity of the node making the proposal. They then confirm whether the node is authorized to make the proposed update. Finally, they confirm whether the proposed update is consistent with the last agreed-upon state of the ledger.
3. *Agreement:* If the update is validated, a majority of nodes will then agree to it.
4. *Recording:* The proposed data modification is permanently recorded in the ledger.[232]

[226] Mills et al., 2016, p. 11.

[227] According to the CPMI, administrative functions may include access control, notarization, dispute resolution, standard setting, and regulatory reporting. See Mills et al., 2016, p. 12; CPMI, 2017, p. 5.

[228] Mills et al., 2016, p. 11; IIF, 2017a, p. 22.

[229] CPMI, 2017, p. 4.

[230] Mills et al., 2016, p. 13.

[231] CPMI, 2017, p. 4.

[232] CPMI, 2017, pp. 4–5.

A DLT arrangement's consensus mechanism partly depends on its type of access control. Permissionless arrangements use consensus mechanisms that generate an economic incentive for nodes to participate in the validation process.[233] Permissioned systems do not depend on economic incentives, since they can assign this task to particular nodes.[234] Permissioned arrangements use a wider variety of consensus mechanisms.[235]

A distributed ledger may display the *last agreed-upon state* or it may display the entire history of *state changes*. For example, a DLT arrangement that facilitates financial transactions may either maintain a record of current ownership or retain the entire history of transactions.[236]

The term *blockchain* refers to the way in which data is structured and stored on a particular type of DLT, such that the ledger retains the entire history of data modifications. On a blockchain, every update must first be agreed to by a majority of the nodes on the network. Once the update has been agreed to, it is combined with other contemporaneous updates into a time-stamped batch or "block." This block is then cryptographically linked to the previous block, making a chain.[237] Each time a new block is added, the entire chain updates simultaneously across the network, such that every node always retains a complete and up-to-date record of all data modifications that have ever taken place.

Advantages of DLT and use cases for AML/CFT compliance

DLT's most exciting promise is that it makes secure disintermediation possible. This expands the scope for cooperation and economic activity in domains where trusted intermediaries are expensive, unreliable, or absent.[238] However, DLT has other attributes that may make it useful even where total disintermediation is unnecessary. This is why financial institutions, themselves a type of intermediary, are interested in the technology.

Financial institutions are exploring a number of potential use cases for DLT. These include, for example, applying blockchain solutions to cross-border payments and foreign exchange (see Box 6), to securities and loan settlements, and to derivatives and other contracts.[239] The two attributes that make blockchain attrac-

[233] CPMI, 2017, p. 4. The most common consensus mechanism is the proof-of-work mechanism, in which nodes compete for the right to validate and record new updates by solving a cryptographic puzzle. The winner of the competition receives a prize, typically paid out in the form of virtual currency. The solution to the puzzle is random, which ensures that the competing nodes have an equal chance of winning (abstracting from disparities in computing power). Proof-of-work is highly secure but also inefficient, due to the fact that it is inherently frictional. The main alternative consensus mechanism for permissionless arrangements is proof-of-stake, which also uses a competition but weights the probability of winning by the nodes' stake in the system (for example, the amount of virtual currency they hold). Proof-of-stake is less computationally intensive than proof-of-work but may be less secure. See Pisa and Juden, 2017, pp. 11, 37–39.

[234] CPMI, 2017, p. 4.

[235] These include practical Byzantine fault tolerance and the Stellar consensus protocol. See Pisa and Juden, 2017, p. 4.

[236] Mills et al., 2016, p. 13.

[237] Pisa and Juden, 2017, p. 1.

[238] Pisa and Juden, 2017, pp. 5–6.

[239] IIF, 2015, pp. 3–9.

Box 6. DLT for Cross-Border Payments[a]

In addition to its potential AML/CFT compliance applications, DLT might help address de-risking by making cross-border payments faster, cheaper, and more transparent.

Cross-border payments are often slow and expensive. This is due to the fragmented nature of the global payments system, with its sprawling network of bilateral CBRs.[b]

The inefficiency of cross-border payments hinders economic development. It makes it expensive for migrants to send remittances home and for export-oriented small and medium-sized enterprises (SMEs) to send and receive payments.[c]

Remittances are an important source of development finance, but despite a concerted push by the international development community, remittance fees remain high.[d] The global average cost of remitting US$200 was 7.4 percent in 2016—more than twice the 3.0 percent target established in the Sustainable Development Goals.[e]

Export-oriented SMEs are broadly dissatisfied with cross-border payments. In one survey, more than two-thirds of SME respondents expressed unhappiness with the delays and fees they encounter when sending and receiving such payments.

Several start-ups are exploring ways to use blockchain technology to improve cross-border payments. Their approaches generally fall into one of three categories: (1) using a virtual currency as a bridge between national currencies, (2) introducing a distributed ledger between banks, and (3) using a "connector" to enhance interoperability between banks' existing private ledgers.[f] The third model is currently the most popular, as it works with banks' existing systems and minimizes many of DLT's potential risks.

1. *Using virtual currency as a bridge:* In this model, a virtual currency, such a bitcoin, serves as a vehicle currency for converting the currency of the payment originator into the currency of the payment beneficiary. Since the international leg of the transaction takes place on the bitcoin network, this method bypasses the correspondent banking system altogether. This is the approach taken by virtual currency–based money transfer operators such as BitPesa, Rebit.ph, and Veem. However, this model has drawbacks: swapping into and out of bitcoin incurs its own transaction fees, and in some corridors, these fees are expensive enough to negate some or all of the value of the model. Partly for this reason, many of the start-ups in this space have either closed or transitioned to different business models.
2. *Using a distributed ledger between banks:* In this model, the underlying payments architecture used by correspondent banks is replaced with distributed ledgers. Banks using these platforms can transact in any currency, with virtual currencies serving as a bridge when the preferred currencies of the payment originator and the beneficiary differ.[g] In such circumstances, the platforms search for the best exchange rates offered by market makers on the network.
3. *Interledger approach:* The Interledger Protocol, developed by Ripple, allows banks to synchronize transactions between their existing private ledgers, rather than requiring them to use the same ledger. This approach improves the speed and transparency of cross-border payments while avoiding the concerns about data privacy, governance, and resiliency that a wholesale switch to DLT might provoke.

When paired with digital identities, these approaches have the potential not only to lower transaction costs and accelerate settlement times, but also to introduce greater transparency and traceability to cross-border payments.

[a] Adapted from Pisa and Juden, 2017, pp. 16–22, with permission.

[b] Pisa and Juden, 2017, p. 16.

[c] Pisa and Juden, 2017, p. 16.

[d] Remittances account for roughly three times the value of official development assistance. See World Bank, 2017b, p. 2.

[e] World Bank, 2017b, p. 4.

[f] Pisa and Juden, 2017, p. 17.

[g] Ripple and Stellar, two start-ups that have developed models along these lines, use their own virtual currencies—XRP and lumen, respectively—for these conversions.

tive for regulatory compliance are that it is tamper-resistant, which improves data integrity, and that it allows for enhanced, nearly instantaneous information sharing. Financial institutions are exploring DLT applications for individual banks as well as for interbank consortia and market infrastructures.

Permissioned DLT arrangements may be used for more secure record keeping. DLT enhances data integrity in two ways. First, once data is recorded on a distributed ledger, it is essentially permanent. The consensus mechanism protects data from tampering or deletion, intentional or otherwise. Moreover, the distributed nature of the ledger ensures that even if some nodes are hacked or otherwise compromised, the remaining nodes—and their copies of the database—will remain unaffected. Second, the fact that the ledger retains the entire history of state changes means that all data modifications are traceable. Auditors of financial institutions could benefit from the immutability and traceability of data stored on distributed ledgers.[240]

Permissioned DLT arrangements could also be used for secure, instantaneous, and reliable information sharing, while nullifying the need for data reconciliation. Since the ledger updates automatically and simultaneously for all nodes on the network, participants can observe activity on the ledger in real time.[241] Moreover, sharing information on a distributed ledger eliminates the need for ledger participants to reconcile their records—a task to which banks currently devote significant resources and one that opens up the possibility for errors.[242]

DLT could facilitate information sharing for AML/CFT compliance, both among banks and between banks and their regulators. Banks and shared utilities could establish closed, permissioned blockchain arrangements for the purpose of sharing KYC information. Regulators could be granted read-only access to audit the ledger.

DLT might be especially useful for sharing KYC information on natural persons. Established KYC utilities are run on centralized databases, so for them, DLT may not necessarily add much value.[243] However, these utilities focus on institutional KYC and thus do not address the parallel issue of costly and redundant KYC checks on natural persons. Privacy laws may preclude the development of central repositories for storing and sharing KYC information on natural persons. DLT-based solutions might be able to address this issue by giving individuals custody of their personal information, along with the power to decide who can access it (see Box 7).[244]

DLT could be used to usher forth deeper, more fundamental changes in AML/CFT compliance. If, someday, all financial transactions were moved onto a distributed ledger, with every transaction linked to the unique digital IDs of the parties involved, it would be possible for banks and government authorities to monitor the entire financial system for illegal activity in real time.[245] Bank supervisors could be granted read-only access to the network to monitor anonymized transaction flow in real time, along with a "master key" to decrypt and investigate suspicious

[240] IIF, 2015, p. 5

[241] GAO, 2017, p. 44.

[242] CPMI, 2017, p. 13.

[243] Pisa and Juden, 2017, p. 21.

[244] Pisa and Juden, 2017, p. 21.

[245] Pisa and Juden, 2017, p. 22.

Box 7. Blockchain for Digital Identification and Personal Information Management[a]

Proponents of using blockchain as a platform for digital ID argue that the technology could underpin the development of "user-centric" or "self-sovereign" ID systems. Such systems could give individuals a secure way to identify themselves and share their personal information online, for purposes including access to financial services, without having to rely on a central database.[b]

In such a system, an individual would store personal information in an "identity wallet" on his or her mobile phone. The wallet would hold documents certified by trusted authorities confirming that the person was who he or she claimed to be and possessed certain attributes.[c] Pisa and Juden (2017) offer an example: "Alice could store the following certified claims in her wallet: 'credit rating over 700,' certified by a bank or credit rating agency; 'has a US passport," or 'is over 21,' certified by the government; 'has blood type B' certified by a hospital or doctor."[d] When the individual wished to grant a third party access to a particular piece of information, he or she could do so without sharing the other information stored in the wallet.[e]

Using blockchain technology to help individuals manage and share their personal information online might confer several benefits. Pisa and Juden (2017) explain: "The first is privacy: Alice can control both who she shares her personal information with and how much information she shares. The second is security, as the absence of a centralized database eliminates single point of failure risk. The system is also more convenient, since it allows users to provide verified information with the touch of a button rather than having to locate and submit a wide variety of documents. Finally, a blockchain provides an easy and accurate way to trace the evolution of ID attributes since each change is time-stamped and appended to the record preceding it."[f]

However, while the benefits of the "user-centric" ID system are obvious in theory, the model's viability depends on stakeholder buy-in. In particular, government authorities must support the system for it to be effective. Moreover, network effects depend on there being a critical mass of customers and financial institutions participating in the system. "If those services only satisfy a small portion of a person's needs," warn Pisa and Juden, "which is likely to be the case if the authorities . . . do not participate, then the value of a user-controlled ID is limited."[g]

[a] Adapted from Pisa and Juden, 2017, pp. 22–27, with permission.

[b] Pisa and Juden, 2017, p. 25.

[c] Lewis, 2017.

[d] Pisa and Juden, 2017, pp. 25–26.

[e] Lewis, 2017.

[f] Pisa and Juden, 2017, p. 26.

[g] Pisa and Juden, 2017, p. 27.

transactions for which they had a subpoena.[246] "The data captured and shared via blockchain allows analysis well beyond the immediate transaction and allows targeted insights into global illicit financial streams," explain Juan Zarate and Chip Poncy, two former US Treasury officials and co-founders of the Financial Integrity Network. When DLT is combined with other technologies, such as big data applications and machine learning algorithms, "the potential for the identification of suspicious patterns and networks increases exponentially."[247]

[246] Pisa and Juden, 2017, p. 22.

[247] Zarate and Poncy, 2016.

Challenges and limitations

Although DLT could potentially be transformative for regulatory compliance, it also faces several challenges. Some of these challenges are well understood, but because DLT is not yet a mature technology, it might possess other shortcomings still waiting to be discovered.

Data privacy may be an obstacle to sharing certain types of information on a distributed ledger. Many DLT applications, including those being considered for regulatory compliance, would link sensitive data to individual identities.[248] This is most obviously true for KYC and transaction data. Permissioned arrangements would have to restrict access to such information by encrypting it, so that participants could read only the information that was directly relevant to them.[249] Even with such precautions, however, it might be possible for participants to make inferences about encrypted identities based on certain patterns or markers in the data.[250] Partly for this reason, financial institutions are wary of putting transaction data on distributed ledgers.[251] In addition, the immutability of DLT may conflict with "right-to-be-forgotten" laws.[252]

Security and operational resiliency could be an issue. The distributed nature of DLT eliminates "single-point-of-failure" risk. However, this does not mean the technology is invulnerable. For example, under a "51 percent attack" scenario, if bad actors were to seize control of a majority of the nodes on a network, they could corrupt the ledger.[253] While this is unlikely to happen on a permissionless ledger that uses a proof-of-work consensus mechanism, closed, permissioned ledgers, which typically have fewer nodes, may be more vulnerable to such an attack. Further, although cryptography protects DLT today, it is possible that future developments in quantum computing will undermine the strength of these protections.[254]

It is possible that more operational vulnerabilities are still waiting to be discovered. As noted by the US Financial Stability Oversight Council, "market participants have limited experience working with distributed ledger systems, and it is possible that operational vulnerabilities associated with such systems may not become apparent until they are deployed at scale."[255]

Effective governance is an inherent challenge for DLT arrangements, due to their peer-driven consensus process. This can make it difficult for DLT arrangements to adjust their protocols to changing circumstances, such as new security risks.[256] Governance challenges are most acute for permissionless arrangements,

[248] Pisa and Juden, 2017, p. 12.

[249] CPMI, 2017, p. 18.

[250] Natarajan, Krause, and Gradstein, 2017, p. 20.

[251] Pisa and Juden, 2017, p. 13.

[252] Pisa and Juden, 2017, p. 13.

[253] Pisa and Juden, 2017, p. 8.

[254] Natarajan, Krause, and Gradstein, 2017, p. 18.

[255] FSOC, 2016, p. 127.

[256] Mills et al., 2016, p. 51.

Table 10. Summary: Advantages and Challenges/Limitations of DLT

Advantages	Challenges
■ Negates need for trust and makes secure disintermediation possible ■ Immutable, transparent, synchronous, and traceable ■ Enables real-time information sharing	■ Privacy ■ Resiliency ■ Governance ■ Interoperability ■ Switching costs ■ Regulatory uncertainty

but they can still be a challenge for permissioned arrangements, too.[257] In consortium arrangements, participants would have proportionately more say as to how the ledger is governed, but none would have total say. Permissioned arrangements can address these governance challenges by differentiating participants' roles and functions on the network, including centralizing certain governance decisions with an administrator.[258] However, the ability of the administrator to enforce its decisions on all network participants will depend on the arrangement.[259]

Interoperability is an important determinant of the scalability of DLT applications. For consortia arrangements to deliver on their promise of greater efficiency, they must achieve scalability. However, at the moment, many different competing DLT applications are being developed, with no clear indication of which, if any, will come to dominate the market.[260] In the absence of collaboration on standardized protocols, it is possible that no single solution or group of solutions will dominate, thus preventing the promised scalability.

Finally, it is possible that DLT's advantages are exaggerated and do not outweigh the costs of transitioning to a new data architecture. It is unclear whether DLT would lead to major improvements in KYC compliance. For institutional KYC, established KYC utilities use centralized databases, not distributed ones, and do not appear to suffer for it. For individual KYC, some countries are developing e-KYC arrangements that give financial institutions access to centralized government databases. Finally, DLT's superior operational resiliency, compared with centralized databases, may be overstated, since the latter can always be backed up.[261]

[257] The inability to resolve protocol disputes can lead to "forking," a situation in which the ledger splits into competing arrangements. See Pisa and Juden, 2017, pp. 13–14.

[258] Mills et al., 2016, p. 31.

[259] Natarajan, Krause, and Gradstein, 2017, pp. 18–19.

[260] IIF, 2015, pp. 9, 13.

[261] Pisa and Juden, 2017, p. 13.

Table 11. Recommendations

Organizations Involved	Recommendation
National regulators	Consider how to amend data sharing and privacy laws to enable sharing of identification, due diligence, and transaction data between banks or between banks and authorities.
Blockchain participants	Begin working to establish common interoperability standards to ensure that different arrangements can integrate with each other.

Prospects for adoption

DLT is still an emerging technology. In a report to the United Kingdom's Financial Conduct Authority (FCA), PA Consulting Group observes that "adoption remains at an almost universally early stage."[262] Banks are exploring various DLT applications through proofs-of-concept and pilot projects, but few, if any, of these are ready for release. The United States' Federal Reserve Board judges that practical blockchain applications are still years away.[263]

Several established financial institutions are experimenting with DLT applications for sharing KYC information. In Singapore, for example, several banks have partnered with the Info-communications Media Development Authority, a government agency, to develop a proof of concept for KYC on the blockchain.[264]

Companies are also introducing new platforms based on the "identity wallet" model described in Box 7. KYC-Chain and Tradle are two start-ups that have developed platforms that allow customers to record KYC verifications in a "digital wallet," which they can share with other financial institutions upon request.[265] Deloitte Luxembourg, a consultancy, has created a similar proof of concept called KYCstart (pronounced "kick-start"), which it says gives customers control over their own data.[266] In Canada, SecureKey has proposed a blockchain for identity verification for online services.[267]

However, private-sector stakeholders disagree on how useful DLT applications will be for AML/CFT compliance. In discussing the results of a survey of banks, start-ups, and technology vendors, PA Consulting Group reported that "blockchain was by far the most contentious of the new technologies explored [for AML compliance] during the course of this study. Opinions varied significantly across all types of respondents, with some considering it unimpressive while others believed it was the 'solution' to AML compliance."[268] Those who were unimpressed considered blockchain a solution in search of a problem, even as they acknowledged the technology's transformative potential. "This unusual juxtaposition (that people can

[262] PA Consulting Group, 2017, p. 24.

[263] Mills et al., 2016, p. 3.

[264] Mui, 2017.

[265] See KYC-Chain here: https://kyc-chain.com/; see Tradle here: https://tradle.io/.

[266] Deloitte, 2017.

[267] See SecureKey here: https://securekey.com.

[268] PA Consulting Group, 2017, p. 22.

think it both potentially incredibly useful, but not think of any specific cases) was repeated by a number of different respondents," PA observed. "The most common view espoused (from both technology and regulated firms) was that truly compelling blockchain use cases had yet to be articulated in AML compliance, restricting the pace of adoption."[269]

Banks' own technical capabilities—and those of regulators—may also limit the pace of adoption. In the same report, PA noted that many established financial institutions felt they lacked the technical expertise to adopt DLT-based AML compliance solutions. They perceived the FCA's expertise in this area to be similarly lacking. "Firms appear reluctant to build a solution that the regulator might be unable to consider or approve," PA observed.[270]

For these applications to be viable, regulators will need to weigh in on the issues of reliance on third parties and of data privacy. The FSB notes that DLT applications may not conform with laws in all jurisdictions, which would limit their cross-border applicability.[271] This is especially relevant with respect to data privacy and localization laws. DLT's structure may also make it difficult to identify who is responsible "for an asset when no one bank is the custodian of the record."[272] In addition, as with KYC utilities, consortia for sharing KYC information on the DLT would be dependent on regulators' permission for banks to rely on third parties for due diligence.[273]

[269] PA Consulting Group, 2017, pp. 15 and 22.

[270] PA Consulting Group, 2017, p. 11.

[271] Mills et al., 2016, p. 31.

[272] Mills et al., 2016, p. 19.

[273] He et al., 2017, p. 24.

DLT: Distributed Ledger Technology

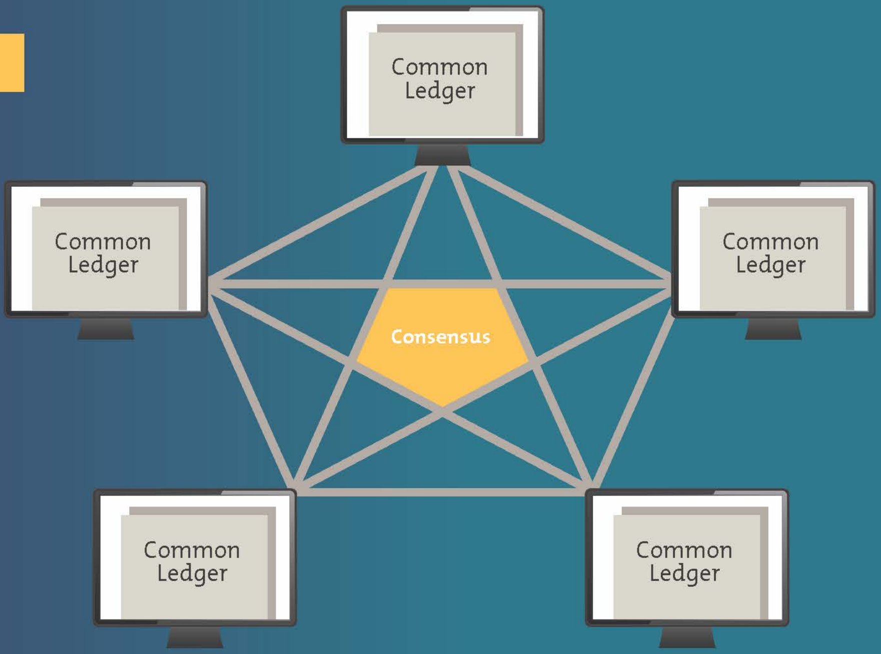
Distributed ledger technology (DLT) is used to securely manage data on a peer-to-peer network of computers.

DISTRIBUTED LEDGER

Synchronized Data Sharing

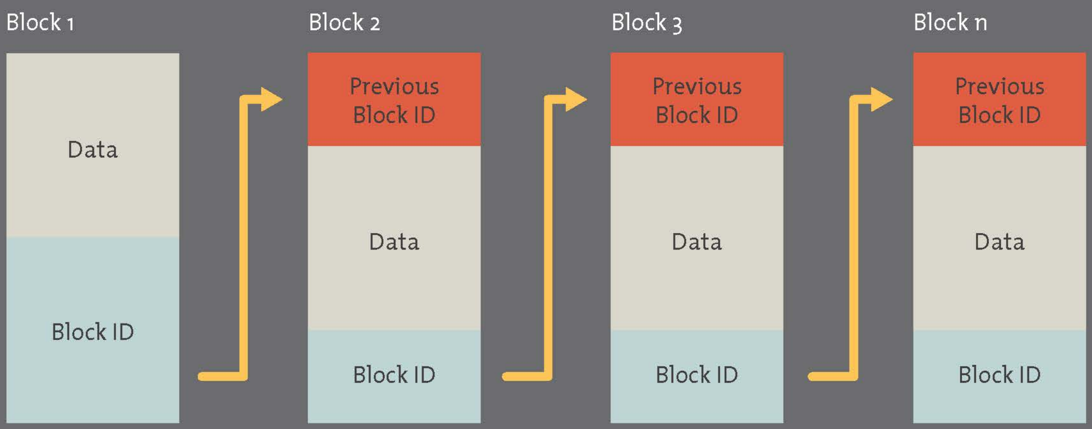
Consensus-Based Governance

Cryptography

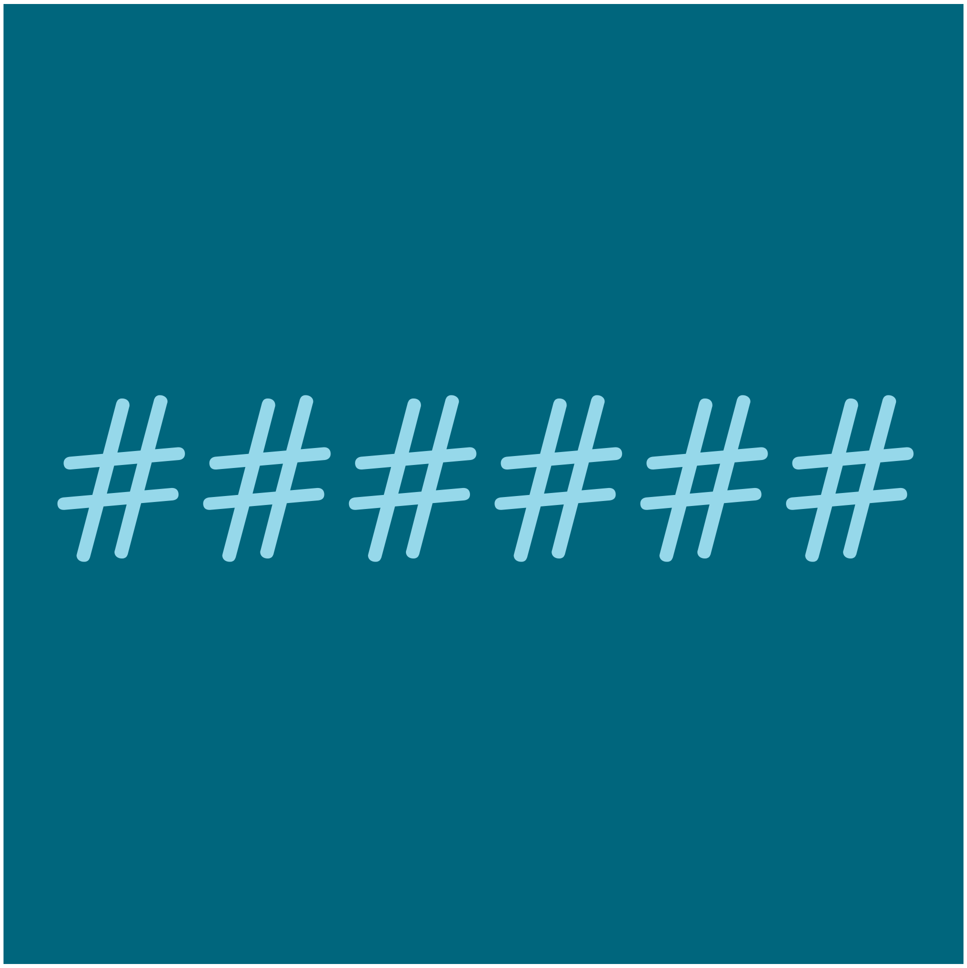


BLOCKCHAIN

Blockchain is a type of DLT in which data modifications are recorded in time-stamped "blocks," each of which is connected to the previous block, forming a chain.



6 IDENTITY MANAGEMENT FOR LEGAL ENTITIES: LEGAL ENTITY IDENTIFIERS



Key points

- LEIs are unique alphanumeric reference codes used to precisely identify legal entities.
- LEIs connect to a reference dataset that contains a standardized set of information about the entity.
- All legal entities—financial institutions, nonfinancial corporations, government agencies, and nonprofit organizations (NPOs)—are eligible to apply for LEIs.
- Natural persons are not eligible to apply for LEIs, except when they are operating in a business capacity—for example, as a sole trader.
- The Global LEI System (GLEIS) is a federated system. It is overseen by the LEI Regulatory Oversight Committee (LEI ROC), administered by the Global LEI Foundation (GLEIF), and implemented by Local Operating Units (LOUs, also referred to as LEI Operating Units).
- LEIs may be used to clearly identify respondent banks and to serve as a starting point for CDD.
- LEIs can support interoperability with other AML applications for conducting KYC on respondent banks by serving as a common reference identifier for KYC utilities and information-sharing arrangements.
- In the future, LEIs could be used to identify payment originators and beneficiaries, thus facilitating easier and more accurate KYCC.

Box 8. The 2015 Center for Global Development Report on LEIs

The 2015 Center for Global Development Working Group on the Unintended Consequences of Anti–Money Laundering Policies identified challenges in client identification as one driver of compliance costs: "The expected costs for a given transaction . . . increase in the probability that a client has not been identified successfully." This problem is worse, for both individuals and institutions, in developing countries.[a] The working group continued, "even in advanced economies, there is not yet a universally accepted unique identifier for institutions."[b]

In the report's recommendations, the Working Group encouraged banks to accelerate their adoption of the LEI:

Recommendation 5: Facilitate Identification and Lower the Costs of Compliance: Legal Entity Identifiers: Banks and other financial institutions should accelerate the global adoption of the Legal Entity Identifier scheme . . . The FSB and national regulators should also enable the adoption of LEIs.[c]

[a] CGD Working Group, 2015, p. 52.

[b] CGD Working Group, 2015, p. 52.

[c] CGD Working Group, 2015, p. 55.

What LEIs are

LEIs are unique 20-character alphanumeric codes assigned to legal entities that engage in financial transactions or enter into contracts. They are often likened to barcodes for legal entities. Their purpose is to precisely identify parties in financial transactions.

All legal entities that enter into financial transactions or contracts are eligible for LEIs. These include not only financial institutions and nonfinancial corporations, but also investment funds, government agencies, and international institutions.[274]

The LEI standard explicitly excludes natural persons, except when they are serving in an independent business capacity.[275] Eligible natural persons may include, for example, CEOs, directors, major traders, and others whose roles allow them to act on behalf of a legal entity.[276]

The LEI was developed in response to the global financial crisis. Its original purpose was to enable banks and regulators to aggregate risk data quickly and accurately (see Box 9). Most recently, LEIs have been used to reduce the opacity of financial transactions involving holding companies.

The LEI was designed to meet a number of specifications. As outlined by the LEI Trade Association Group and, separately, by Bottega and Powell, the specified features included:

- *Singular (exclusive):* There is only one LEI per entity.
- *Unique:* LEIs should never be reassigned, reused, or shared.
- *Persistent:* LEIs should stay with the entity regardless of name changes, location changes, and so on.
- *Neutral:* The LEI number is "dumb." The number itself is random and does not include any embedded information, such as data to indicate the company's name, industry, or location (which can change).
- *Extensible (scalable):* The LEI number should be long enough to accommodate significant growth in the number of identifiers being used.
- *Structurally fixed:* The LEI format must not change over time.
- *Reliable:* The LEI issuance process is intended to ensure that the data is of high quality.
- *Interoperable:* The LEI should be machine readable and not conflict with other ID systems.
- *Publicly available:* The LEI database should be free to access and download. There should be no restrictions on use.[277]

[274] GLEIF, n.d.–c; Interview with Stephan Wolf, CEO, GLEIF, August 23, 2017.

[275] Kennickell, 2016, p. 11.

[276] A natural person's information must be verifiable in a public data source, such as a company filing (interview with Stephan Wolf, August 23, 2017).

[277] LEI Trade Association Group, 2011, pp. 19–20; Bottega and Powell, 2011, pp. 10–12.

Box 9. A Short History of the LEI

Before the global financial crisis, there was no common, standardized identification system for legal entities. Banks, vendors, and government agencies all had their own incompatible identification systems. In many cases, banks' internal ID systems were further split along organizational or functional lines, due to internal competition, the absence of firmwide data governance strategies, and mergers and acquisitions.[a] It was a financial "Tower of Babel," as Andrew Haldane put it in a 2012 speech.[b]

Previous attempts to develop a single identification standard were unsuccessful. In the two decades preceding the global financial crisis, industry attempts to unite around a single identification standard for legal entities foundered in the face of "competing priorities, funding issues, and lack of industry focus (among many other issues)."[c]

Attempts to cross-reference different identification schemes likewise fell short. These attempts to map identifiers between banks and between banks and government agencies were hampered by "ambiguities and inconsistencies in those relationships [that] often [made] cross-referencing difficult and inaccurate."[d] For a bank to aggregate data across business units or for a government agency to aggregate data across the entities it oversaw was a time-consuming, manual process of matching up identifiers, often with the business name being the only common link between datasets.[e]

Within banks, the use of various incompatible ID systems made data aggregation difficult and error-prone. Analyzing risk data required staff to first manually consolidate it, which could take anywhere from days to weeks.[f]

As a result, banks had trouble monitoring their firm-wide risk exposures to various counterparties. Moreover, regulators could not discern the connections between firms, which made it impossible to forecast risk contagion.[g]

The global financial crisis revealed the dangers posed by these disparate and incompatible identification systems. When Lehman Brothers collapsed, its counterparties could not easily assess their exposure to the investment bank. Regulators could not predict how the failure would propagate through the financial system.[h] The result was panic.

Shortly thereafter, policymakers and industry groups came together to design a common identification standard for financial institutions. In 2010, the US Department of the Treasury's Office of Financial Research issued a policy statement calling for the creation of a global LEI.[i] The following May, a coalition of financial-sector industry groups responded with their requirements. Shortly thereafter, the FSB assumed responsibility for the LEI's development.

ISO issued Standard 17442 in May 2012. The G20 endorsed the LEI two months later.

GLEIS is underpinned by a charter, not a legally binding treaty. During negotiations on the form of GLEIS, it was decided that the treaty process was too cumbersome to be practical. Instead, GLEIS was agreed to by charter, with charter members agreeing to incorporate the LEI into domestic laws and regulations as they saw fit.[j]

[a] Senior Supervisors Group, 2010, p. 10.

[b] Haldane, 2012, p. 11.

[c] Bottega and Powell, 2011, p. 3.

[d] Bottega and Powell, 2011, p. 4.

[e] Bottega and Powell, 2011, p. 5.

[f] Senior Supervisors Group, 2010, p. 10.

[g] This fragmentation stood in stark contrast to the product markets, which adopted the Uniform Product Code (the bar code) in the 1970s, the Global Location Number in the 1980s, and the Serial Shipping Container Code in the 1990s. These unique identifiers provided businesses with a clear and comprehensive view of their supply chains, which further enabled a host of innovations, including supply-chain automation and just-in-time manufacturing. See Haldane, 2012, p. 4.

[h] According to Grody and Hughes, "There was no consistency in identifying Lehman as a counterparty. No understanding of what relationships Lehman had with others; no mechanism to associate all of Lehman's products and businesses into a total view of the exposure others had to Lehman should it fail. In effect, no one—not regulators nor creditors nor counterparties—could see into Lehman's exposure to risk" (2015, pp. 19–20).

[i] Office of Financial Research, US Department of the Treasury, 2010.

[j] Couillault, Mizuguchi, and Reed, 2017, p. 8.

The LEI may be used to look up the entity's reference data, or business card, which includes a standardized set of information. These reference data are specified in the Common Data File format, which is subject to occasional revision. "Level 1" data—described as the "who is who" data[278]—include basic data, such as:

1. Official name
2. Country of legal formation
3. Headquarters address
4. Dates of LEI issuance, most recent updates, and (if applicable), expiration

"Level 2" data—described as the "who owns whom" data—includes information on parent entities, where applicable.[279] This information is publicly available. The database may be accessed on the GLEIF website.[280] In the future, the LEI may also include the Entity Legal Form code, another ISO standard, published in July 2017, that is used to identify an entity's legal form, such as limited liability partnership, S-corporation, and so on.

How GLEIS works

GLEIS is a federated identity system. LEIs are issued by a limited number of approved entities. These entities operate within a unified governance structure. Federated identity systems may be contrasted with centralized identity systems (in which identities are issued by a single entity) and distributed identify systems (in which many entities issue identification).[281]

GLEIS is governed by LEI ROC. LEI ROC's members include representatives from more than 70 regulatory authorities. It is overseen by an executive committee that is intended to be regionally representative.

GLEIF is the central operating unit. Incorporated in Switzerland and based in Germany, it is responsible for overseeing GLEIS and ensuring that the LOUs adhere to LEI standards and principles. It is also responsible for promoting the LEI and for identifying new use cases, as appropriate.

The LOUs are responsible for implementation. Their responsibilities include registration, data validation, records maintenance, publication, and periodic data checks. There were 30 LOUs as of mid-2016. The largest is the Global Market Entity Identifier GMEI Utility, jointly run by SWIFT and the DTCC.

When a legal entity registers with an LOU, it is assigned a unique LEI. This LEI will never be assigned to any other entity. Moreover, the legal entity in question can never obtain another LEI, either in addition to the one it first received or as a substitute.

LEIs are issued by LOUs for a fee. The system is operated as a not-for-profit entity on a cost-recovery basis. Fees are nominal.

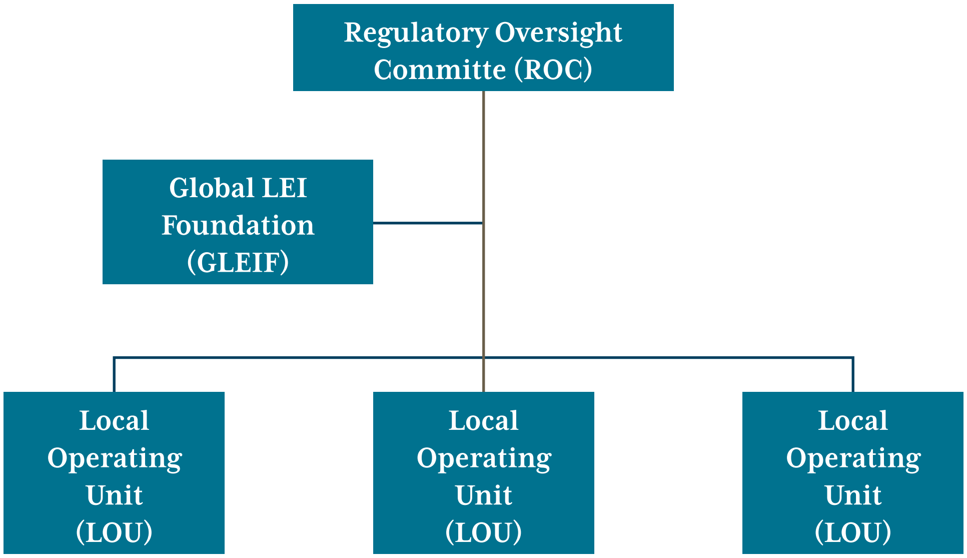
[278] GLEIF, n.d.–a.

[279] GLEIF, n.d.–a.

[280] https://www.gleif.org.

[281] WEF, 2016, p. 53.

Figure 1. Global LEI System Governance Structure


Regulatory Oversight Committe (ROC)
Global LEI Foundation (GLEIF)
Local Operating Unit (LOU)
Local Operating Unit (LOU)
Local Operating Unit (LOU)

The LOU is responsible for verifying the information submitted by the applying legal entity or its registration agent. Where possible, the LOU is expected to verify against official sources, such as business registries. If official sources are unavailable or unreliable, the LOU may use "authoritative private sources," such as an official record of incorporation.[282] The LOU includes the level of validation (i.e., "fully corroborated," "partially corroborated," "entity supplied only," or "pending") and the sources of validation in the LEI's reference data.[283] LOUs are expected to resolve any discrepancies that arise. Moreover, users may file a challenge if they believe the reference data contains erroneous or outdated information, which the LOU must then investigate.[284]

As of January 2018, more than a million LEIs have been issued.[285]

Advantages of the LEI and use cases for AML/CFT compliance

Although the LEI was not originally intended for either payments or AML/CFT compliance, it can be adopted for these purposes. The LEI was originally intended for micro- and macroprudential risk management, but its design is sufficiently flexible that it can serve as a general-purpose reference identifier. Its creators assumed that it would find new use cases over time.

[282] Kennickell, 2016, p. 14.

[283] LEI ROC, 2014, p. 26.

[284] PMPG, 2016, p. 8.

[285] GLEIF, n.d.–b.

The current ID system for cross-border transactions does not allow for precise identification of legal entities in all cases. In cross-border transactions for both financial and nonfinancial institutions, payment messages currently rely on BICs for account identification and routing destination.[286] Correspondent banking IT systems are built around the BIC.[287] However, BICs are not ideal for unambiguous identification. Some legal entities may use more than one BIC, and in a few cases, a single BIC can be used by multiple entities within a group to send or receive transaction messages.[288]

LEIs are superior to BICs for identifying transacting parties precisely, especially for risk management and compliance purposes, where error tolerance is low. While an entity may have more than one BIC, a legal entity may have only one LEI.

LEI reference data can serve as a starting point for conducting CDD on respondent banks. When two companies enter into a relationship, information is exchanged about who they are. The number of exchanges raises the possibility that there will be an error. Instead of receiving name, address, and other basic data from the onboarding customer, which runs the risk of transcription errors, the correspondent bank could simply receive the onboarding customer's LEI and then use the LEI to look up the respondent bank's basic information.[289]

LEIs can serve as the foundation for other AML applications for conducting KYC on respondent banks. Standardized reference identifiers represent a common language. They make it easier to automate processes and to share data across systems and institutions.[290] Large financial institutions have begun using deep learning to map their identifiers across institutions, but this practice may not provide the level of accuracy needed for AML/CFT compliance.[291] LEIs can be used as reference identifiers for a range of applications. LEIs would make it easier for financial institutions to identify which transacting parties are already customers of theirs, reducing unnecessary CDD checks. They could be used in automated sanctions screening programs to reduce the rate of false positives (in cases in which names and/or addresses partly match the data).[292] They could also be used with KYC utilities. Although SWIFT uses the BIC for its KYC Registry (as mentioned above), other KYC utilities use different identifiers. This makes mapping between different KYC utilities (to aggregate information on a single entity among them) very difficult.[293]

In the future, LEIs could be used to identify payment originators and beneficiaries, thus facilitating easier and more accurate KYCC. The CPMI has noted that the LEI offers one route to fulfilling FATF recommendation 16, which calls for the inclusion of information on originators and beneficiaries in payments messages.[294]

[286] SWIFT, n.d.–b.

[287] ABA, 2015, p. 3.

[288] SWIFT, 2015, p. 4.

[289] Interview with Paul Janssens, LEI programme director, SWIFT, August 21, 2017.

[290] SWIFT and Deloitte, 2012.

[291] Interview with Stephan Wolf, August 23, 2017.

[292] CPMI, 2016, p. 24.

[293] SWIFT, 2015, p. 4.

[294] CPMI, 2016, p. 25.

Such a capability would enable better screening of payments messages, which could be done either in real time or through batch processing.[295] However, this capability does not yet exist (more on that below).[296] The use of LEIs in payment messages would also make it easier for financial intelligence units to consolidate transaction information reported by different financial institutions, thereby enabling effective systemwide surveillance.[297]

Challenges and limitations

The LEI was not designed for KYC. That is, although the LEI can serve as a starting point for customer identification, verification, and due diligence, it is not a substitute. Banks may not rely on the information included in a legal entity's reference data, but must instead verify it on their own.

The LEI reference dataset does not include all of the information necessary to conduct CDD—most importantly, CDD related to beneficial ownership. Although GLEIS has begun collecting information on the parents of legal entities, it does not collect information on the *beneficial owners* of legal entities, who are natural persons. This is, in part, because the LEI is a public database, and many countries do not allow information on beneficial owners to be made public.[298] In addition, the LEI reference data does not include industry classification codes, which are useful for customer typology, behavioral analysis, and predictive analytics.

Currently, LEIs are used mainly by financial institutions. If LEIs are to be used in payments messages to identify originators and beneficiaries, they will have to be more widely adopted by nonfinancial corporations.[299] However, awareness of the LEI is not widespread among nonfinancial corporations, and so adoption is slow.[300] Financial institutions do not believe LEI adoption can be driven by the banking sector alone: "for the LEI to become an effective tool it needs to be embedded in commercial transactions and hence originators of payments need to capture the LEI of their counterparties in [enterprise resource planning] systems."[301]

The LEI does not apply to natural persons—a big gap in AML/CFT. If LEIs were to be used in payments messages to identify originators and beneficiaries, they would face an additional limitation, which is that they would not cover individuals engaging in transactions. As the ABA has noted, the fact that LEIs do not apply to natural persons "precludes a large portion of the transactions from the LEI application."[302]

[295] Interview with Paul Janssens, August 21, 2017.

[296] ABA, 2015, p. 3.

[297] CPMI, 2016, p. 24.

[298] Interview with Paul Janssens, August 21, 2017; interview with Stephan Wolf, August 23, 2017.

[299] PMPG, 2017, p. 2.

[300] PMPG, 2016, p. 2.

[301] PMPG, 2017, p. 2.

[302] ABA, 2016, p. 2.

Legacy payment message formats—MT103 and MT202 COV—do not include a field for the LEI and are not expected to introduce one. The LEI could be included in a free-form field, though. Discussions are underway regarding whether to include an LEI field in the new ISO 20022 standard for payment messages.

The costs associated with updating bank IT systems to incorporate the LEI may, in the short term, discourage widespread adoption in payment messaging systems.[303] Although it is cheap for legal entities to register for an LEI, it is expensive for banks to update their payment messaging systems, which are built around the BIC.[304] This is made more challenging by the fact that the benefits of adopting the LEI would likely be diffused across the bank, but the costs would have to be borne by a particular business unit, which may have difficulty shouldering or justifying the expense on its own.[305] This could prove to be a significant hurdle in the short term. In a Payments Market Practice Group (PMPG) survey of financial institutions, respondents expressed concerns about the costs of adoption, which they anticipated would be high.[306] They suggested that if the LEI is to be required in payment messages, regulators should wait for the adoption of the new ISO 20022 standard. Use of the LEI in legacy payment message formats should be voluntary.[307]

Some financial institutions consider the information currently included in payment messages generally sufficient to perform CDD. In the PMPG survey, some respondents said that they did not believe there was a compelling case for incorporating the LEI into payments messages, arguing "that the LEI does not address any existing problems and that current practices work well."[308] In the PMPG survey, several respondents indicated that "the BIC was sufficient for identifying banks in the payment chain."[309]

New technical capabilities will be required to incorporate the LEI into transaction monitoring systems for real-time screening of originators and beneficiaries. The Clearing House and the IIB argue that banks do not have the ability to use LEIs to screen for KYCC information, writing, "functionalities do not yet exist to enable banks to obtain updated information on a respondent bank's customers based on an analysis of payment flows—even when payment messages include ISO country codes, the LEI, or other information." The Clearing House and the IIB further point out that the ability to analyze the flow of messages of the respondent bank to provide information on the customers using correspondent banking services "is not yet available, and . . . would only be possible if KYC utilities have direct access to payment message information."[310]

[303] PMPG, 2016, p. 12.

[304] Interview with Paul Janssens, August 21, 2017.

[305] Interview with Paul Janssens, August 21, 2017.

[306] PMPG, 2017, p. 2.

[307] PMPG, 2017, pp. 2–3.

[308] PMPG, 2017, p. 1.

[309] PMPG, 2017, p. 1.

[310] The Clearing House and IIB, 2017, pp. 12, 14.

Prospects for adoption

The LEI is now mandatory for certain financial transactions in the United States, and will be soon for nearly all financial transactions in the European Union, but it is less widely used in developing countries. The US Commodity Futures Trading Commission became the first regulator to mandate the LEI's use, in this case for derivatives counterparties, in 2012. The EU mandated the use of the LEI for derivatives counterparties shortly thereafter; it then mandated the use of LEIs for money market transactions in 2016. Beginning in 2018, EU law will require LEIs for investment firms and counterparties to securities transactions. This will effectively render the LEI a de facto identifier for most financial market participants in the EU, as well as outside of it (per extraterritoriality), as EU reporting firms must identify external counterparties by their LEIs as well.[311]

Stakeholders are discussing whether and how best to incorporate the LEI into current and future payment message formats. For legacy payment message formats (MT103 and MT202 COV), most stakeholders favor allowing banks to include the LEI as an optional value in the free-form field. Stakeholders are discussing whether to include an LEI field in the new ISO 20022 payment messages standard.

More needs to be done to drive LEI adoption, particularly for developing countries and nonfinancial corporations. Regulatory requirements have been a major driver of LEI registration in the United States and in the European Union. However, fewer developing countries mandate the LEI's use, and so the pace of adoption is lower.[312] One exception is Mexico, whose central bank now requires all depository institutions (and their counterparties) to obtain an LEI, which is then linked to their tax ID number.[313] Other developing-country signatories to the LEI charter should look for ways to incorporate the LEI into their regulations, as appropriate.

LEI adoption may also be driven by the private sector. Although regulatory mandates have been an important driver of LEI registration in the first few years, financial institutions and nonfinancial corporations can also encourage or require their counterparties to register for LEIs. In addition, legal entities may independently decide that it is worthwhile to obtain an LEI. Where it is feasible, such organic growth may be preferable to regulatory mandates. Since it depends on legal entities' recognizing the LEI's value for themselves, it may be more sustainable.[314]

In addition to low adoption rates, developing countries are also often hindered by lower-quality information. Business registries may contain partial or outdated information. This means LOUs have a greater challenge in verifying the data that is submitted, lengthening the registration process.[315]

[311] Interview with Paul Janssens, August 21, 2017.

[312] Interview with Paul Janssens, August 21, 2017.

[313] LEI ROC, 2015, p. 13; Annex I, p. 11; interview with Stephan Wolf, August 23, 2017.

[314] Interview with Paul Janssens, August 21, 2017.

[315] Interview with Paul Janssens, August 21, 2017.

Table 12. Summary: Advantages and Challenges/Limitations of LEI

Advantages	Disadvantages
■ A strong mechanism for identifying transacting parties involved in payment chains, to support KYC and KYCC ■ Reduces fiction/uncertainty as well as processing times to identify counterparties—increases certainty of identification ■ Can improve reliability of information by providing essential background information ■ Makes data aggregation easier ■ Has a strict data validation process ■ Competition among LOUs incentivizes a low price for LEI registration ■ Can facilitate automation and interaction between institutions and platforms ■ Can reduce false positives in client screening by reducing misidentification ■ Could free employees from work of matching and tagging	■ Not a substitute for banks' due diligence ■ Banks cannot rely on information—must independently verify ■ Cannot be used to identify natural persons in payments chain (except in particular circumstances) ■ Does not help with beneficial ownership ■ Adoption in payments system may be expensive ■ Dependent on network effects

Some industry groups believe that the LEI's effect on de-risking would be limited. The ABA thinks adopting LEIs for cross-border payments may increase efficiency in the long run but "in itself will not provide relief to correspondent banks making the decision to discontinue ongoing correspondent banking relationships due to rational concerns about regulatory risk." The ABA recommends that regulators provide guidance on how banks can rely on LEIs for compliance purposes.[316]

What regulators, policymakers, and standard-setting bodies are doing to facilitate the appropriate adoption of the LEI for AML/CFT

Stakeholders are taking steps to increase LEI registrations. To promote LEIs among nonfinancial institutions, GLEIF began a campaign in April 2017, encouraging financing institutions to become "registration agents" of LEI issuers.[317] This would allow banks to facilitate LEI applications on behalf of their customers. SWIFT recommends that all entities be required to create an LEI when joining a KYC utility.[318]

[316] ABA 2015, p. 3.

[317] FSB, 2017b, p. 17.

[318] SWIFT, 2015, p. 4.

Table 13. Recommendations

Organizations Involved	Recommendation
Standard-setting bodies	Determine whether LEIs can be used for customer identification, verification, and due diligence, and provide relevant guidance.
National regulators in countries affected by de-risking	Look for ways to promote LEI issuance.
National regulators in countries affected by de-risking	Improve business registries and other relevant information sources that LOUs use to validate information.
Financial institutions	Help customers obtain LEIs, especially in countries affected by de-risking.
Banks	Begin modifying IT systems to prepare for adoption of LEIs in payment messages.
ISO	Continue work on how best to incorporate the LEI into the new payment messaging format.

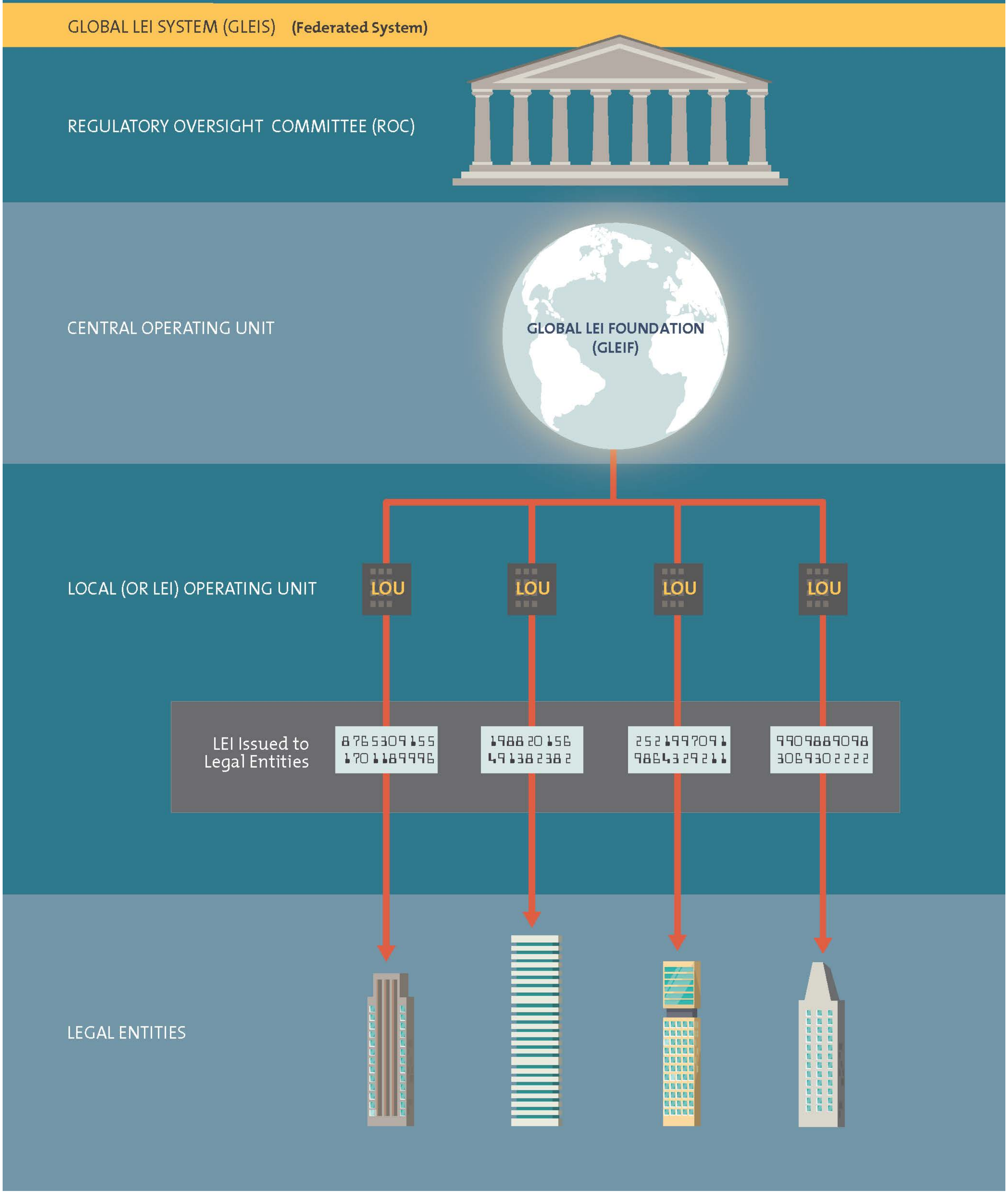
GLEIF and SWIFT are developing a BIC-to-LEI mapping facility. This tool will enable banks to utilize the LEI without necessitating a change in current payment message formats; it will also mean that banks will not have to conduct the mapping themselves—they can simply download the mapping table.[319] SWIFT is designated by ISO as the registration authority for BICs.

The CPMI has recommended allowing LEIs to be included in payment messages on a voluntary basis. The LEI would be entered in the free-form field, without altering the payment message's format or standard data fields.

[319] Interview with Stephan Wolf, August 23, 2017.

Legal Entity Identifiers

LEIs are unique 20-character alphanumeric codes assigned to legal entities that engage in financial transactions or enter into contracts. They are often likened to barcodes for legal entities. Their purpose is to precisely identify parties in financial transactions.



GLOBAL LEI SYSTEM (GLEIS) **(Federated System)**

REGULATORY OVERSIGHT COMMITTEE (ROC)

CENTRAL OPERATING UNIT

LOCAL (OR LEI) OPERATING UNIT

LEI Issued to Legal Entities

LEGAL ENTITIES

7 IDENTITY MANAGEMENT FOR NATURAL PERSONS: BIOMETRICS

Key points

- Because LEIs do not apply to natural persons, a separate standard is needed for identifying individuals engaging in financial transactions.
- Biometrics use distinctive physiological or behavioral characteristics to authenticate a person's identity and control his or her access to a system.
- Biometric identifiers cannot help correspondent banks with KYCC, as they are unlikely to be included in payment messages any time in the foreseeable future. However, the use of biometrics by respondent banks, money transfer organizations (MTOs), and NPOs may enhance correspondent banks' confidence in the AML/CFT controls of these counterparties.
- Biometrics may be used to address the "identification gap" that exists in many developing countries. This, in turn, could make it easier for banks to conduct customer identification, verification, and due diligence.
- Compared with other "authentication factors," such as passwords and tokens, biometrics are more secure and easier to use.
- A number of biometric identification systems have been developed in recent years at the national level. Work is needed to develop an internationally recognized or interoperable digital identification system.

Box 10. The 2015 Center for Global Development Report on Biometrics

The Center for Global Development Working Group on the Unintended Consequences of Anti–Money Laundering Policies (2015) identified challenges in client identification as one driver of compliance costs. With respect to the problem of identifying individuals, the working group argued that "national identification systems sufficient for customer identification" were within reach for "the vast majority of countries," citing India's Aadhaar program (which assigns unique 12-digit identifiers to Indian residents based on biometric data) along with "weaker identification systems" in 37 other countries.[a]

In the report's recommendations, the working group encouraged national governments in developing countries to provide their residents with identification robust enough to be used for KYC purposes:

> **Recommendation 5: Facilitate Identification and Lower the Costs of Compliance:** Identification of Individuals: National governments should provide citizens with the means to identify themselves in order to make reliably identifying clients possible for financial institutions and other organizations. . . . [Further,] national governments should ensure that appropriate privacy frameworks and accountability measures support these identification efforts while ensuring the free flow of information related to identifying ML [money laundering] and TF [terrorist financing].[b]

[a] CGD Working Group, 2015, p. 53.

[b] CGD Working Group, 2015, p. 55.

What biometrics are

Digital biometric recognition systems use distinctive physiological or behavioral traits to automatically identify and authenticate individuals.[320] Biometrics are a technical solution to three related but distinct functions:

- *Identification:* Who are you?
- *Authentication:* Are you who you claim to be?
- *Authorization:* What rights and privileges do you have in this system?[321]

A number of distinctive physiological and behavioral characteristics can be used for biometric identification. Distinctive physiological characteristics include fingerprints; hand, finger, or palm geometry; vein patterns in the finger or back of the hand; facial features; facial thermograms; iris patterns; blood vessel patterns in the retina; and DNA.[322] The three most commonly used traits are fingerprints, irises, and facial features.[323] Distinctive behavioral characteristics that can be used for biometric identification include keystroke dynamics, signature dynamics, and voiceprints.[324]

Biometrics fall into one of three broad categories of authentication mechanisms, commonly referred to as authentication factors:[325]

- *Knowledge factors:* Something only you know, such as a password, a personal identification number (PIN), or your mother's maiden name.
- *Possession (aka object or token) factors:* Something only you have, such as a driver's license, a key, or an ATM card.
- *Inherence factors:* Something only you are (biometrics).[326]

With the broad adoption of GPS-enabled devices in recent years, some have suggested considering geolocation—somewhere you are—as a fourth authentication factor.[327]

The accuracy and security of biometric systems can be enhanced through the use of *multi-layered* or *multifactor* authentication. Mult-layered biometric authentication is the use of more than one biometric trait, such as a fingerprint and a voiceprint. Multifactor authentication is the use of more than one authentication factor, such as biometrics in combination with knowledge-based or token-based credentials.[328] For example, when individuals are using GPS-enabled devices, biometric systems can be combined with geolocation data.

The choice of which trait to use depends on several factors. Jain, Ross, and Nandakumar (2011) outline seven factors that should generally be considered:

[320] Jain, Nandakumar, and Ross, 2016, p. 80.

[321] Gelb and Clark, 2013, p. 1; Lott, 2015, p. 3.

[322] Lott, 2015, pp. 21-27.

[323] Jain, Nandakumar, and Ross, 2016, p. 82.

[324] Lott, 2015, pp. 27–29.

[325] Lott, 2015, pp. 5, 12–17.

[326] O'Gorman, 2003, p. 2024.

[327] Lott, 2015, p. 5.

[328] Lott, 2015, pp. 8–9.

1. *Universality:* Every individual accessing the application should possess the trait.
2. *Uniqueness:* The given trait should be sufficiently different across individuals making up the user population.
3. *Permanence:* The biometric trait of an individual should be sufficiently invariant over a period of time with respect to the matching algorithm.
4. *Measurability:* It should be possible to acquire and digitize the trait using suitable devices that do not cause undue inconvenience to the individual. Furthermore, the acquired raw data should be amenable to processing in order to extract discriminative feature sets.
5. *Performance:* Apart from recognition accuracy, the computational resources required to achieve that accuracy and the throughput (number of transactions that can be processed per unit of time) of the biometric system should also meet the constraints imposed by the system.
6. *Acceptability:* Individuals in the target population that will utilize the application should be willing to present their biometric trait to the system.
7. *Circumvention:* This refers to the ease with which the trait of an individual can be imitated using artifacts (e.g., fake fingers), in the cases of physical traits, and mimicry, in the case of behavioral traits. It also refers to obfuscation, where a user deliberately alters his biometric trait to evade recognition.[329]

Other analytical frameworks have been suggested for judging the suitability of biometric systems as they relate specifically to payments and mobile financial services. Lott suggests a six-factor framework comprising robustness (permanence), distinctiveness, accessibility, availability, acceptability, and financial (cost).[330] Lovisotto and colleagues use survey data to derive a five-factor framework (with several subfactors), comprising modality performance, usability, interoperability, security, and privacy.[331]

What identity systems are

Identity systems (whether biometric-based or not) can be either *foundational* or *functional*.[332] Foundational identity systems, such as national ID systems, are *supply driven*—that is, their purpose is to supply a general-purpose formal ID, which can be used for both public and private interactions.[333] In contrast, functional identity systems are *demand driven*—that is, their purpose is to facilitate access to some activity or service, such as driving, voting, or receiving transfer payments.[334] Functional IDs can sometimes serve as de facto foundational IDs, as Social Security numbers do in the United States.[335]

[329] Jain, Ross, and Nandakumar, 2011, pp. 29–30.

[330] Lott, 2015, p. 20.

[331] Lovisotto et al., 2017, pp. 6–12.

[332] Gelb and Clark, 2013, p. 3.

[333] Gelb and Clark, 2013, pp. 20, 35–38.

[334] Gelb and Clark, 2013, pp. 20, 23–35.

[335] Gelb and Clark, 2013, p. 20.

ID systems can be developed by national governments, but they can also be developed by industry consortia or by individual institutions. This is particularly true of functional identity systems.

Biometrics were once a costly advanced technology, but not anymore. In recent years, the price of technology to capture and authenticate biometric features has fallen dramatically.

How biometric authentication works

Biometric systems work in two stages: enrollment and recognition.[336] During enrollment, the biometric system uses a sensor to capture the enrollee's biometric trait(s) for the first time. The system then obtains the trait's distinctive features and uses that *feature set* to generate a *template*—a mathematical representation—which is then (usually) encrypted.[337] Often, the system then compares the new template with the other templates stored in its database to ensure uniqueness. This process is known as *identification* or *de-duplication* and is used to counter fraud by ensuring that no one is enrolling in the system twice.[338] De-duplication is important for voter ID and social security systems.[339] Finally, the system generates a unique identifier, such as a reference number, which connects the feature set to the user.[340] Once the trait(s) have been captured and stored in the system, they can be used as an authenticator.

During recognition, the enrolled user presents his or her biometric trait to access the system. Once again, the individual presents him- or herself to a sensor, which then captures the user's biometric traits again and obtains the distinctive features. If multifactor authentication is being used—for example, if the user enters an ID number or swipes a card—the system can then directly retrieve the user's template stored in the database and compare the new feature set with that of the original template.[341] Alternately, if no other form of identification is being provided, the system may check the template against all other templates stored in the system to find the match. The original template may be stored on a central database, if the reader and the database are connected to the Internet, or it may be stored on a card.[342]

One important way in which biometric systems differ from knowledge-based authentication systems is that the latter require an exact match, whereas the former do not. If, for example, the password being entered does not exactly match the password stored in the database, the system will not authenticate the user.[343] With a biometric system, an exact match is not always possible. Differences may arise between sample template and the stored template, due to any of a number of factors, such as problems or differences with the sensors, sweat, dirt, low light,

[336] Jain, Nandakumar, and Ross, 2016, p. 81.

[337] Jain, Nandakumar, and Ross, 2016, pp. 81–82.

[338] Gelb and Clark, 2013, p. 63.

[339] Lott, 2015, p. 17.

[340] Mas and Porteous, 2012, p. 50.

[341] Lott, 2015, p. 17.

[342] Lott, 2015, p. 17.

[343] Lott, 2015, p. 19.

glare, and so on.[344] With a biometric system, an algorithm is used to determine the probability that the match is correct.[345]

What problems biometrics addresses

Biometric identifiers cannot help correspondent banks with KYCC. Payments messages do not include biometric identifiers, and they are unlikely to do so for the foreseeable future. Accordingly, biometric identifiers cannot help correspondent banks to more accurately identify payments' originators and beneficiaries.

That said, the use of biometrics by respondent banks, MTOs, and NPOs may enhance correspondent banks' confidence in the AML/CFT controls of these counterparties. When correspondent banks onboard new clients—be they respondent banks, MTOs, or charities—they must conduct due diligence on them. This risk assessment includes an assessment of the prospective client's own AML/CFT controls.[346] One of the drivers of de-risking is that correspondent banks often lack confidence in their clients' risk management capabilities.[347]

For respondent banks' risk controls to be effective, it is critical for them to know who their customers are—both during the onboarding phase and in subsequent transactions. Banks are required to ascertain the identity of their customers using "reliable, independent source documents, data, or information."[348] Reliable customer identification processes are the foundation of all subsequent due diligence.

However, the ability to positively identify customers can be a challenge in poor countries, which often suffer from "identification gaps." Some developing countries lack formal identification systems that could be used to verify customers' identities.[349] Even those that do have such systems often suffer from coverage gaps. Poor people are especially likely to lack formal identification.

The lack of "formal identification" makes customer identification, verification, and due diligence a challenge for financial institutions in many developing countries. For example, the IMF has reported that money transfer operators in the Pacific Islands struggle to perform CDD because their countries do not have national identification systems in place.[350]

Further, customer identification and due diligence are often paper-based in developing countries. This practice may reduce correspondent banks' confidence in the respondent banks' ability to manage due diligence information or to respond to information requests.[351]

Digital identities based on biometric authenticators may be used to address two problems for respondent banks, MTOs, and charities: identification for CDD and authentication for transaction monitoring.

[344] Gelb and Clark, 2013, p. 64.

[345] Gelb and Clark, 2013, p. 5.

[346] The Clearing House, 2017, p. 10.

[347] FSB, 2017a, p. 45; IMF, 2017, p. 1.

[348] FATF, 2016b, p. 14.

[349] Gelb and Clark, 2013, pp. 7–8.

[350] IMF, 2017, p. 2.

[351] IMF, 2017, p. 30.

Identity for onboarding and CDD

Biometrics may be used to address the identity gap that exists in many developing countries. Biometrics offer the most effective way for developing countries to establish national identity systems.[352] Such initiatives may be driven by national governments, by individual banks, or by industry consortia.

In turn, identity systems can make it easier for banks to conduct KYC. Biometrics can be used to uniquely identify people and to authenticate them. If a customer's ID is linked to a data repository containing basic information about the customer, that can be used to conduct due diligence.

In countries where there is a national ID, banks and money transfer operators can use that credential to onboard new customers. For example, in Malaysia, each citizen is issued a smart card with biometric information, called myKad. Some banks, such as RHB Banking Group, allow customers to open new accounts using myKad readers.[353]

Authentication for secure transaction authorization

Biometrics may be used for secure transaction authorization. Banks are often challenged to verify the identities of transacting customers with a high degree of certainty, especially when these customers are remote. As banks manage more and more transactions remotely, there is a growing need to ensure that transaction authorizations are secure.

Passwords and PINs are still the most widely used authentication factors for financial transactions. This is because they are low cost, do not need to be carried, and can be revoked or changed if they are compromised.[354]

However, passwords and PINs are insecure and increasingly difficult for customers to manage. First, they are vulnerable to hacking. Most passwords are easy to crack. "Based on an all-or-nothing approach," explains Accenture, "they afford no protection once they have been compromised."[355] Digital fraud and identity theft have been rising at an alarming pace.

Passwords and PINs are also difficult to remember. With the proliferation of online services, users have increasing difficulty managing passwords. Customers often forget their PINs or passwords and therefore must reset them, at great cost to financial institutions—as much as 30 percent of all call center calls are for password resets.[356] Password managers may be employed, but they are also vulnerable to hacking.[357]

[352] Gelb and Clark, 2013, p. 12.

[353] Accenture, 2013, p. 12.

[354] Lovisotto et al., 2017, p. 1

[355] Accenture, 2013, p. 6.

[356] Accenture, 2013, p. 6.

[357] Lovisotto et al., 2017, p. 1.

Biometrics offer greater security and greater ease of use than do passwords or PINs. Biometrics are among the most secure and robust authenticators. Biometric fraud requires the fraudster to obtain the user's biometric trait and, in the case of mobile devices, the device on which the user enrolled.[358] In addition, unlike knowledge- or token-based authenticators, biometrics cannot be lost or forgotten.

Challenges and limitations

As stated above, biometrics are not being considered for inclusion in international payments messages and thus do not lend additional transparency to correspondent banking transactions.

Although their security is robust, biometrics are not invulnerable to fraud. Fingerprints can be faked, for example.

Prospects for adoption

Biometrics are now an affordable, mature, and widely used technology. According to a report by PA Consulting Group to the UK FCA, the use of biometrics in AML and KYC has become commonplace in recent years and that biometrics are now regarded "one of the most mature and instantly useful elements of technology in AML."[359]

Fingerprints and facial recognition are the most commonly used biometric identifiers, as these traits work well with mobile phones.[360] Voice recognition is also becoming widely adopted in call centers.[361]

Legal identification is now considered an important development goal. The UN's Sustainable Development Goals (SDGs), adopted in 2015, include a target—SDG Target 16.9—aimed at "providing legal identity to all, including birth registration" by 2030.[362]

Many developing countries have embarked on biometric-based identification programs in recent years. Worldwide, there are 37 national biometric ID programs, most of which rely on fingerprints.[363] The most prominent is India's national ID system, Aadhaar. Launched in 2010, it is now the largest biometric-based ID system in the world, having succeeded in its goal of enrolling nearly all of India's adult residents.[364] Residents who enroll are assigned a unique 12-digit number; their identity is authenticated by fingerprints and iris scans. Aadhaar is run by the Unique Identification Authority of India, part of the Ministry of Electronics and Information Technology.

[358] Lovisotto et al., 2017, p. 3.

[359] PA Consulting Group, 2017, p. 19.

[360] Lovisotto et al., 2017, p. 3.

[361] Lott, 2015, p. 28.

[362] United Nations, 2015.

[363] UN ITU, 2016, p. 8.

[364] Prasad, 2017.

Table 14. Summary: Advantages and Challenges/Limitations of Biometrics

Advantages	Challenges
▪ By addressing the identification gap in developing countries, can help make it easier for banks to conduct customer identification, verification, and due diligence ▪ Can be used to identify individuals in transactions ▪ Security ▪ Ease of use ▪ Portability	▪ Not invulnerable to fraud ▪ Not useful for KYCC in correspondent banking (for the foreseeable future)

Several supranational initiatives are devoted to expanding formal identification in developing countries. One such initiative is the World Bank's Identification for Development (ID4D) program, launched in 2014. Another is ID2020, a public-private partnership launched in 2016.[365] In June 2017, Microsoft, Accenture, and Avanade introduced a prototype digital identity network that utilizes biometric and blockchain technologies.[366]

To date, the only program that has considered using biometrics specifically to address de-risking is the Safer Corridor Pilot—a plan to ensure the continued flow of remittances from the United Kingdom to Somalia. This plan was devised by the UK's Department for International Development (DFID), FSD Africa (a DFID-funded nonprofit organization), the World Bank, and Consult Hyperion (a consultancy that specializes in payments technology).[367] However, the plan was never executed, due in part to the fact that remittance flows to Somalia continued.[368]

Assessments by regulators, policymakers, standard-setting bodies, and trade associations

The FSB believes that biometric authentication may be a solution to problems with remittances. It is considering establishing a work stream in this area. Similarly, the IMF has noted that "improving identity verification systems within a country may be warranted" in certain cases.[369]

The FATF has pointed approvingly to the use of biometrics. In its guidance on reconciling AML/CFT with financial inclusion, the FATF highlighted "innovative technological solutions" for customer identification and verification, including public and private biometric solutions in Malawi, India, and New Zealand.[370]

[365] http://id2020.org.

[366] Irrera, 2017.

[367] Consult Hyperion and FSD Africa, 2017, p. 9.

[368] Erbenová et al., 2016, p. 35.

[369] Erbenová et al., 2016, p. 38.

[370] FATF and World Bank, 2013, pp. 78–79.

What regulators, policymakers, and standard-setting bodies can do to facilitate the appropriate adoption of biometrics for AML/CFT

Despite the proliferation of ID systems (biometric and otherwise), they remain fragmented, typically along jurisdictional lines—financial institutions have no common way of identifying individuals across jurisdictions.[371] In recent years, the number of biometric identity schemes worldwide has grown substantially. However, these systems are fragmented, having been created at the national or subnational level. Such systems lack interoperability. They have "different codes [and] different allocation rules, and in many cases [are] not even unique."[372]

Policymakers should do more to explore what is needed to develop an internationally recognized, interoperable digital identification system for natural persons. As remote interactions increase, especially across borders, there is increasingly a need to consider developing an internationally recognized digital identification system for individuals. Such a system would dramatically reduce financial institutions' KYC costs.[373]

The implementation of GLEIS has highlighted the demand for an individual identifier. Stephan Wolf, CEO of GLEIF and co-convener of ISO's Fintech Technical Advisory Group, observed that "progress in [LEI] implementation exposed a need to identify certain types of individuals acting in a business capacity." This was not something for which the LEI was originally intended, and so work-arounds had to be developed—for example, inputting individuals' first and last names in the "legal entity name" field.[374] Wolf writes that "there appears to be sufficient justification to consider a new ISO standard for identifying natural persons, for multiple public and private purposes."[375]

ISO is considering how best to move forward on the development of an internationally recognized identification standard for natural persons. This work is in its initial stages. ISO is currently determining whether it could adapt an existing ISO standard for this purpose, or whether it would have to develop a new standard.[376] "Substantial discussion would be required," writes Wolf, "to determine whether such a standard is appropriate, what are the necessary elements for identification, who should serve as the registrar(s) for the identification, how the data should be managed, and the terms under which it could be made available."[377]

A global system for the identification of natural persons could borrow certain design elements from the LEI but would likely differ in other respects. Such an identifier could use a code structure similar to the LEI's.[378] It could also

[371] Wolf, 2017, p. 4.

[372] Wolf, 2017, p. 5.

[373] Wolf, 2017, p. 7.

[374] Wolf, 2017, p. 3.

[375] Wolf, 2017, p. 4.

[376] Wolf, 2017, p. 5.

[377] Wolf, 2017, p. 4.

[378] Wolf, 2017, p. 5.

Table 15. Recommendations

Organizations Involved	Recommendation
Standard-setting bodies	Explore what steps are needed to develop an internationally recognized, interoperable digital identification system for natural persons.
National regulators	Continue to develop biometric-based national ID systems with robust privacy controls.

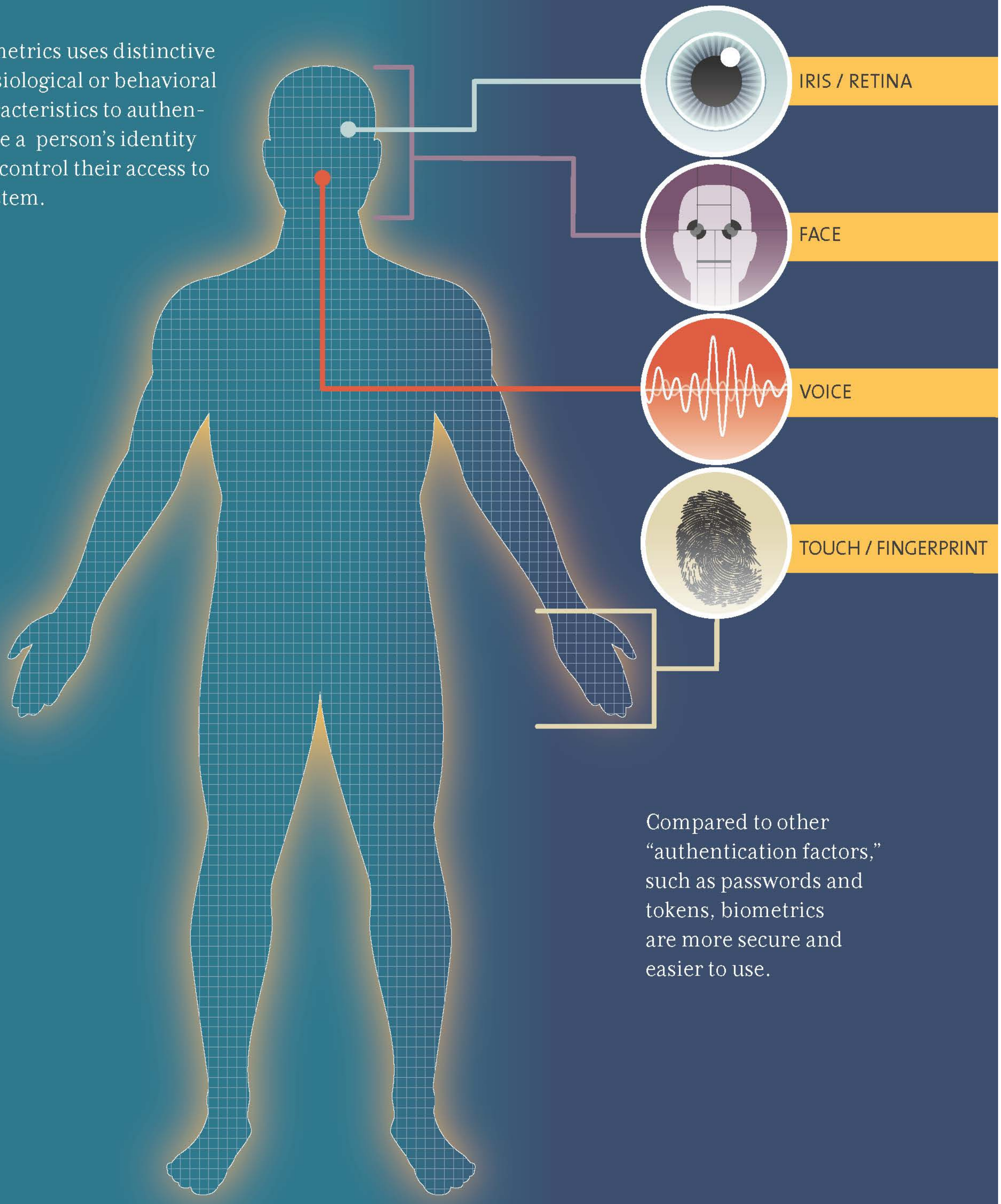
adopt a federated system, similar to that of GLEIS—for example, a "Global Natural Persons Identifier System" (GNPIS) run by a "Global Natural Persons Identifier Foundation."[379] However, a natural-persons identifier would require a different reference dataset than that required by the LEI. In addition, the database would not be public, as it is for the LEI. A GNPIS would work best if it were made to be interoperable with GLEIS. The ability to cross-link identifiers across the two systems would enable relationships between individuals and businesses to be accurately recorded.[380]

[379] Wolf, 2017, p. 9.

[380] Wolf, 2017, pp. 8–9.

Biometrics

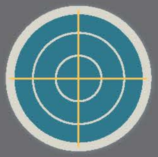
Biometrics uses distinctive physiological or behavioral characteristics to authenticate a person's identity and control their access to a system.



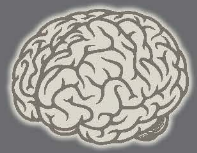
Compared to other "authentication factors," such as passwords and tokens, biometrics are more secure and easier to use.

Can be enhanced by other authentication factors, including:

Location

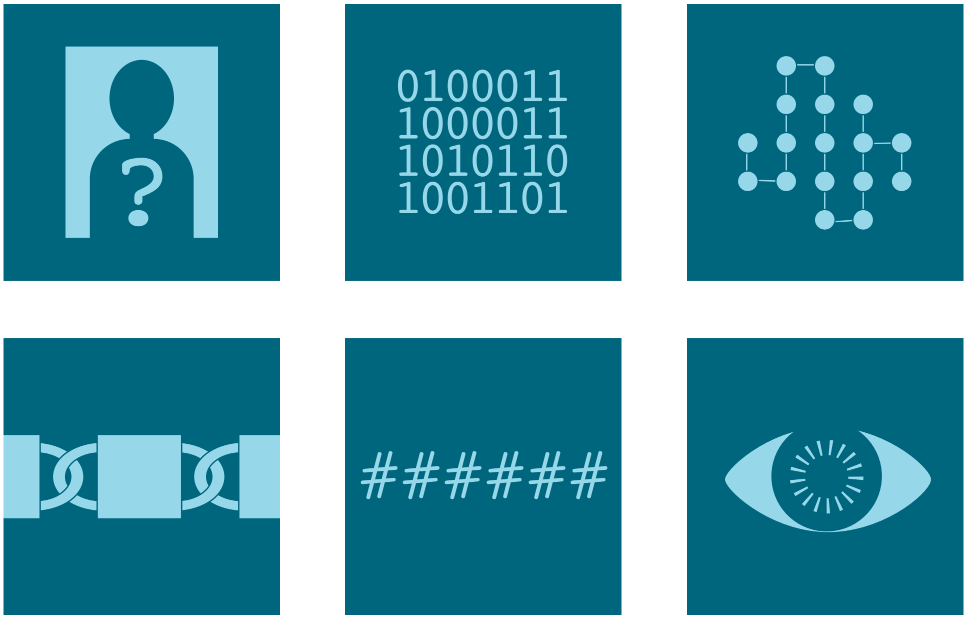


Knowledge Based Factors



Tokens/Passwords

8 CONCLUSION



The technologies described in this report have the potential to significantly improve the effectiveness and efficiency of correspondent banks' anti–money laundering/countering the financing of terrorism (AML/CFT) controls. They do this through a variety of pathways and by targeting different components of the compliance workflow. KYC utilities reduce the amount of time correspondent banks must spend on duplicative customer due diligence processes. Big data and machine learning enhance correspondent banks' ability to assess and manage risk. Distributed ledger technology can improve the security with which information is stored and shared. legal entity identifiers and biometrics can enable faster and more assured identification of legal entities and individuals.

Over time, these technologies may alleviate some of the pressures on banks. The recent downward trend in correspondent banking relationships has a number of causes, not all of which are related to illicit finance. For this reason, these technologies are not a panacea. That said, the rising cost and complexity of AML/CFT regulations, combined with banks' increased sensitivity to money laundering / terrorist financing risk, is a significant driver. Technologies that enable banks to lower costs or to manage risk more confidently should make it easier for banks to offer correspondent banking services.

Policymakers and regulators should consider how best to responsibly engage with and encourage the adoption of these technologies. It is in policymakers' interest to foster more robust AML/CFT capabilities. Both law enforcement and intelligence rely on banks to detect and report suspected illicit finance activities. It is also in policymakers' interest to see that the compliance burden does not prevent banks from offering these services, which may drive down financial connectivity and inclusion, or push financial activity into less regulated channels.

REFERENCES

ABA (American Bankers Association). 2015. Letter to the Committee on Payments and Market Infrastructures, re CPMI Consultative Report on Correspondent Banking. December 7. Washington, DC.

——. 2016. Letter to the Payments Market Practice Group, re discussion paper, “LEI in the Payments Market.” December 20. Washington, DC.

——. 2017. Letter to the Honorable Steven T. Mnuchin, re request for information on Department of the Treasury regulations that can be eliminated, modified, or streamlined. July 31. Washington, DC.

Accenture. 2013. “The Future of Identity in Banking.” London.

Alleyne, Trevor Serge Coleridge, Jacques Bouhga-Hagbe, Thomas Dowling, Dmitriy Kovtun, Alla Myrvoda, Joel Chiedu Okwuokei, and Jarkko Turunen. 2017. “Loss of Correspondent Banking Relationships in the Caribbean: Trends, Impact, and Policy Options.” Working Paper 17/209. Washington, DC: International Monetary Fund.

Alwazir, Jihad, Fazurin Jamaludin, Dongyeol Lee, Niamh Sheridan, and Patrizia Tumbarello. 2017. “Challenges in Correspondent Banking in the Small States of the Pacific.” Working Paper 17/90. Washington, DC: International Monetary Fund.

Andreesen Horowitz. 2017. “AI Playbook.” Menlo Park, CA.

Aruna, P. 2017. “Bank Negara to Make It Easier for Money Transfer.” *The Star* (Malaysia), September 6.

Ayasdi. 2017. “Anti–Money Laundering Solution Deep Dive.” White paper. Menlo Park, CA.

Bauguess, Scott. 2017. “The Role of Big Data, Machine Learning, and AI in Assessing Risks: A Regulatory Perspective.” Keynote Address at OpRisk North America 2017, New York, June 21.

BCBS (Basel Committee on Banking Supervision). 2017. “Guidelines: Sound Management of Risks Related to Money Laundering and Financing of Terrorism.” Basel, Switzerland: Bank for International Settlements.

Bethencourt, Daniel. 2017. “Despite Successes, AML Professionals Still Wary of Artificial Intelligence.” ACAMS MoneyLaundering.com, June 2.

Blackburn, Fred, Josh Sullivan, Peter Guerra, Angela Zutavern, Steve Escaravage, Ezmerelda Khalil Sager, Steve Mills, and Alex Cosmas. 2015. The Field Guide to Data Science, 2nd ed. Washington, DC: Booz Allen Hamilton.

Bottega, John, and Linda Powell. 2011. “Creating a Linchpin for Financial Data: The Need for a Legal Entity Identifier.” Finance and Economics Discussion Series no. 2011-07. Washington, DC: Board of Governors of the Federal Reserve Board.

Capco and Finextra. 2016. “What Makes Utilities Useful? Report on a Survey by Capco and Finextra” Washington, DC: Capco.

CGD (Center for Global Development) Working Group. 2015. “Unintended Consequences of Anti–Money Laundering Policies for Poor Countries.” Washington, DC: Center for Global Development.

Chemitiganti, Vamsi. 2015a. “Big Data—Banking’s New Weapon in War against Financial Crime (1/2).” Vamsi Talks Tech (blog), October 8.

——. 2015b. “Big Data—Banking’s New Weapon in War against Financial Crime (2/2).” Vamsi Talks Tech (blog), October 13.

——. 2017. “Why Data Silos Are Your Biggest Source of Technical Debt.” Vamsi Talks Tech (blog), May 6.

Chemitiganti, Vamsi, and Joe Gillespie. 2016. “The New Frontier in Anti–Money Laundering.” Webinar. Washington, DC: Booz Allen Hamilton; Santa Clara, CA: Hortonworks. April 21.

Chen, Frank. 2016. “AI, Deep Learning, and Machine Learning: A Primer.” Podcast. Menlo Park, CA: Andreessen Horowitz.

The Clearing House. 2016. “Guiding Principles for Anti-Money Laundering Policies and Procedures in Correspondent Banking.” Washington, DC.

——. 2017. “A New Paradigm: Redesigning the US AML/CFT Framework to Protect National Security and Law Enforcement.” Washington, DC.

The Clearing House and IIB (Institute of International Bankers). 2017. Letter to the Basel Committee on Banking Supervision, re consultative document—*Revised Annex on Correspondent Banking*. February 22. Washington, DC.

Cloudera. 2015. “Industry Brief: Anti–Money Laundering and the Enterprise Data Hub in Financial Services.” Palo Alto.

Cognizant. 2014. “OFAC Name Matching and False-Positive Reduction Techniques.” Cognizant 20-20 Insights. Teaneck, NJ.

Consult Hyperion and FSD Africa. 2017. “Anti–Money Laundering, Know Your Customer, and Curbing the Financing of Terrorism.” Nairobi, Kenya.

Couillault, Betrand, Jun Mizuguchi, and Matthew Reed. 2017. “Collective Action: Toward Solving a Vexing Problem to Build a Global Infrastructure for Financial Information.” OFR Brief 17-01. Washington, DC: Office of Financial Research, US Department of the Treasury.

CPMI (Committee on Payments and Market Infrastructures). 2016. “Correspondent Banking.” Basel, Switzerland: Bank for International Settlements.

——. 2017. “Distributed Ledger Technology in Payment, Clearing and Settlement: An Analytical Framework.” Basel, Switzerland: Bank for International Settlements.

CPMI (Committee on Payments and Market Infrastructures) and World Bank Group. 2016. “Payment Aspects of Financial Inclusion.” Basel, Switzerland: Bank for International Settlements.

Deloitte. 2017. “Deloitte Develops Blockchain Proof-of-Concept to Mutualize KYC Checks.” Press release, May 5. Luxembourg.

Diebold, Francis. 2012. “A Personal Perspective on the Origin(s) and Development of ‘Big Data’: The Phenomenon, the Term, and the Discipline.” Unpublished.

Dow Jones and SWIFT. 2017. “Global Anti–Money Laundering Survey Results 2017.” New York: Dow Jones.

Eckert, Sue, Kay Guinane, and Andrea Hall. 2017. “Financial Access for US Nonprofits.” Washington, DC: Charity and Security Network.

The Economist. 2016. “The Digit Era: Indian Business Prepares to Tap into Aadhaar, a State-Owned Fingerprint Identification System.” December 24.

Einav, Liran, and Jonathan Levin. 2013. “The Data Revolution and Economic Analysis.” NBER Working Paper 19035. Cambridge, MA: National Bureau of Economic Research.

Erbenová, Michael, Yan Liu, Nadim Kyriakos-Saad, Alejandro López-Mejía, Giancarlo Gasha, Emmanuel Mathias, Mohamed Norat, Francisca Fernando, and Yasmin Almeida. 2016. “The Withdrawal of Correspondent Banking Relationships: A Case for Policy Action.” IMF Staff Discussion Note. Washington, DC: International Monetary Fund.

FATF (Financial Action Task Force). 2016a. “FATF Guidance: Correspondent Banking Services.” Paris.

——. 2016b. "International Standards on Combating Money Laundering and the Financing of Terrorism and Proliferation." Paris.

FATF (Financial Action Task Force) and World Bank. 2013. "FATF Guidance: Anti–Money Laundering and Terrorist Financing Measures and Financial Inclusion." Paris.

Finextra. 2017. "MAS to Roll Out National KYC Utility for Singapore." March 24.

FSB (Financial Stability Board). 2012. "A Global Legal Entity Identifier for Financial Markets." Basel, Switzerland.

——. 2017a. "Correspondent Banking Data Report." Basel, Switzerland.

——. 2017b. "FSB Action Plan to Assess and Address the Decline in Correspondent Banking: Progress Report to G20 Summit of July 2017." Basel, Switzerland.

FSOC (Financial Stability Oversight Council). 2016. "2016 Annual Report." Washington, DC.

GAO (Government Accountability Office). 2017. "Financial Technology: Information on Subsectors and Regulatory Oversight." Washington, DC.

Gelb, Alan, and Julia Clark. 2013. "Identification for Development: The Biometrics Revolution." Working Paper 315. Center for Global Development: Washington, DC.

GLEIF (Global Legal Entity Identifier Foundation). n.d.–a "Common Data File Formats." Last accessed January 30, 2018.

GLEIF (Global Legal Entity Identifier Foundation). n.d.–b. "LEI Statistics." Last accessed January 30, 2018.

GLEIF (Global Legal Entity Identifier Foundation). n.d.–c "Questions and Answers." Last accessed January 30, 2018.

Grody, Allan, and Peter Hughes. 2015. "Risk, Data and the Barcodes of Finance." *Journal of Financial Transformation* 45:136–158.

Haldane, Andrew. 2012. "Towards a Common Financial Language." Speech at Securities Industry and Financial Markets Association (SIFMA) symposium, "Building a Global Legal Entity Identifier Framework," New York, March 12.

He, Dong, Ross Leckow, Vikram Haksar, Tommaso Mancini Griffoli, Nigel Jenkinson, Mikari Kashima, Tanai Khiaonarong, Celine Rochon, and Hervé Tourpe. 2017. "Fintech and Financial Services: Initial Considerations." IMF Staff Discussion Note 17/05. Washington, DC: International Monetary Fund.

Hilbert, Martin. 2015. "Quantifying the Data Deluge and the Data Drought." Background note for *World Development Report 2016*. Washington, DC: World Bank.

Hilbert, Martin, and Priscila Lopez. 2012. "How to Measure the World's Technological Capacity to Communicate, Store, and Compute Information, Part I: Results and Scope." *International Journal of Communications*, 6:956–979.

IIF (Institute of International Finance). 2015. "Banking on the Blockchain: Re-engineering the Financial Architecture." Washington, DC.

——. 2016. "Regtech in Financial Services: Solutions for Compliance and Reporting." Washington, DC.

——. 2017a. "Deploying Regtech against Financial Crime." Washington, DC.

——. 2017b. "Financial Crime Information Sharing Survey Report." Washington, DC.

IIF (Institute of International Finance) and BAFT (Bankers Association for Finance and Trade). 2015. Letter to the Committee on Payments and Market Infrastructures, re CPMI consultative report on correspondent banking, December 7.

——. 2017. Letter to the Basel Committee on Banking Supervision, re revised annex for correspondent banking to the BCBS guidelines on the sound management of risks related to money laundering and financing terrorism, February 22.

IMF (International Monetary Fund). 2017. "Recent Trends in Correspondent Banking Relationships—Further Considerations." Policy Paper. Washington, DC.

Irrera, Anna. 2017. "Accenture, Microsoft Team Up on Blockchain-Based Digital Network." Reuters, June 19.

Jain, Anil, Karthik Nandakumar, and Arun Ross. 2016. "50 Years of Biometric Research: Accomplishments, Challenges, and Opportunities." *Pattern Recognition Letters* 79:80–105.

Jain, Anil, Arun Ross, and Karthik Nandakumar. 2011. Introduction to Biometrics. Springer: New York.

Kennickell, Arthur. 2016. "Identity, Identification and Identifiers: The Global Legal Entity Identifier System." Finance and Economics Discussion Series 2016-103. Washington, DC: Board of Governors of the Federal Reserve System.

Klein, Aaron, and Kristopher Readling. 2015. "Acceleration in Suspicious Activity Reporting Warrants Another Look." Bipartisan Policy Center blog, September 15.

KPMG. 2014. "Global Anti–Money Laundering Survey 2014." New York.

Lagarde, Christine. 2017. "Working Together to Fight Money Laundering and Terrorist Financing." Speech at Financial Action Task Force plenary meeting, Valencia, Spain, June 22.

Laney, Doug. 2001. "3D Data Management: Controlling Data Volume, Velocity, and Variety." META Group research note. Stamford, CT: Gartner.

LEI ROC (Legal Entity Identifier Regulatory Oversight Committee). 2014. "LEI Data File Format 1.0." Basel, Switzerland.

——. 2015. "Progress Report by the Legal Entity Identifier Regulatory Oversight Committee (LEI ROC): The Global LEI System and Regulatory Uses of the LEI." Basel, Switzerland.

LEI Trade Association Group. 2011. "Requirements for a Global Legal Entity Identifier (LEI) Solution." Basel, Switzerland: Legal Entity Identifier Regulatory Oversight Committee.

Lewis, Antony. 2017. "A Gentle Introduction to Self-Sovereign Identity." *Bits on Blocks* (blog), May 17.

LexisNexis Risk Solutions. 2016. "Uncover the True Cost of Anti–Money Laundering & KYC Compliance." Irvine, CA.

Loffi, Daniel. 2016. "Independent Outlook on KYC Landscape." Presentation, PricewaterhouseCoopers Risk Services, Hong Kong, August 25.

Lott, David. 2015. "Improving Customer Identification." Retail Payments Risk Forum working paper. Federal Reserve Bank of Atlanta.

Lovisotto, Giulio, Raghav Malik, Ivo Sluganovic, Marc Roeschlin, Paul Trueman, and Ivan Martinovic. 2017. "Mobile Biometrics in Financial Services: A Five Factor Framework." Oxford, UK: Department of Computer Science, University of Oxford.

Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. 2011. "Big Data: The Next Frontier for Innovation, Competition, and Productivity." Washington, DC: McKinsey Global Institute.

Mas, Ignacio, and David Porteous. 2015. "Minding the Gaps." *Innovations* 10 (1–2): 27–52.

Mayo, Daniel. 2016. "Assessing the Role of Big Data in Tackling Financial Crime and Compliance Management." London: Ovum.

Miller, Rena, and Liana Rosen. 2017. "Anti–Money Laundering: An Overview for Congress." Washington, DC: Congressional Research Service.

Mills, David, Kathy Wang, Brendan Malone, Anjana Ravi, Jeff Marquardt, Clinton Chen, Anton Badev, Timothy Brezinski, Linda Fahy, Kimberley Liao, Vanessa Kargenian, Max Ellithorpe, Wendy Ng, and Maria Baird. 2016. "Distributed Ledger Technology in Payments, Clearing, and Settlement." Finance and Economics Discussion Series 2016-095. Washington, DC: Board of Governors of the Federal Reserve System.

Minority Staff of the Permanent Subcommittee on Investigations. 2001. "Correspondent Banking: A Gateway for Money Laundering." Washington, DC: Permanent Subcommittee on Investigations, Committee on Homeland Security & Governmental Affairs, United States Senate.

Mitchell, Tom. 1997. *Machine Learning*. McGraw Hill Education: New York.

Mui, Rachel. 2017. "Singapore Regulator, OCBC, HSBC, MUFG Create 'Know Your Customer' Blockchain Prototype." *The Business Times* (Singapore), October 3.

Natarajan, Harish, Solvej Karla Krause, and Helen Luskin Gradstein. 2017. "Distributed Ledger Technology (DLT) and Blockchain." FinTech Note 1. Washington, DC: World Bank Group.

Nelson, Hector, John Plansky, Vlad Gil, and Danny Ludeman. 2016. "Market Utilities in Financial Services: What Role Will You Play?" New York: PwC.

NextAngles. 2016. "Anti–Money Laundering Survey Results." New York.

NIST (National Institute of Standards and Technology) Big Data Public Working Group. 2015a. "NIST Big Data Interoperability Framework: Volume 1, Definitions." NIST Special Publication 1500-1. Washington, DC: NIST, US Department of Commerce.

———. 2015b. "NIST Big Data Interoperability Framework: Volume 2, Big Data Taxonomies." NIST Special Publication 1500-2. Washington, DC: NIST, US Department of Commerce.

———. 2015c. "NIST Big Data Interoperability Framework: Volume 3, Use Cases and Requirements." NIST Special Publication 1500-3. Washington, DC: NIST, US Department of Commerce.

———. 2015d. "NIST Big Data Interoperability Framework: Volume 6, Reference Architecture." NIST Special Publication 1500-6. Washington, DC: NIST, US Department of Commerce.

Office of Financial Research, US Department of the Treasury. 2010. "Statement on Legal Entity Identification for Financial Contracts." Washington, DC.

O'Gorman, Lawrence. 2003. "Comparing Passwords, Tokens, and Biometrics for User Authentication." Proceedings of the *IEEE* 91 (12): 2021–2040.

Oracle. 2013. "Big Data & Analytics Reference Architecture." Oracle Enterprise Transformation Solutions Series. Redwood Shores, CA.

PA Consulting Group. 2017. "New Technologies and Anti–Money Laundering Compliance." London: United Kingdom Financial Conduct Authority.

Pasquali, Valentina. 2017. "Automated Screening Systems Still Plagued by False Positives." *ACAMS moneylaundering.com*, May 8.

PCAST (President's Council of Advisors on Science and Technology). 2014. "Report to the President: Big Data and Privacy: A Technological Perspective." Washington, DC: Executive Office of the President.

Petrasic, Kevin, Steve Chabinsky, Benjamin Saul, and Kelly Chamberlain. 2017. "Managing Risk Through AI: What Financial Institutions Need to Know." Webinar, Association of Certified Anti–Money Laundering Specialists, July 11.

Petrasic, Kevin, Benjamin Saul, and Matthew Bornfreund. 2017. "The Emergence of AI RegTech Solutions for AML and Sanctions Compliance." *Risk & Compliance*, April–June.

Pisa, Michael and Matt Juden. 2017. "Blockchain and Economic Development: Hype vs. Reality." CGD Policy Paper 107. Washington, DC: Center for Global Development.

PMPG (Payments Market Practices Group). 2016. “LEI in the Payments Market.” La Hulpe, Belgium: Society for Worldwide Interbank Financial Telecommunication (SWIFT).

———. 2017. “Use of the Legal Entity Identifier (LEI) in Payments—Status Update.” La Hulpe, Belgium: Society for Worldwide Interbank Financial Telecommunication (SWIFT).

Prasad, Shri Ravi Shankar. 2017. “UIDAI Achieves 111 Crore Mark on Aadhaar Generation.” Press release, January 27. New Delhi: Unique Identification Authority of India (UIDAI).

Protiviti. 2017. “The Challenges of Managing a Global AML Program.” Menlo Park, CA.

PwC. 2015. “Share and Share Alike: Meeting Compliance Needs Together with a KYC Utility.” London.

Ramachandran, Mukund. 2016. “A Horizontal Solution to a Vertical Problem: Why Segmentation Matters to Banks.” Ayasdi blog. March 8.

Ray, Arin. 2015. “Emergence of Utility Model Case for KYC Solutions.” Webinar, April 2. Celent: New York.

Ray, Arin and Neil Katkov. 2016. “Artificial Intelligence in KYC-AML.” Celent: New York.

Readling, Kristofer. 2016. “Artificial Intelligence and Anti–Money Laundering.” Bipartisan Policy Center blog, July 26.

Samuel, Arthur. 1959. “Some Studies in Machine Learning Using the Game of Checkers.” *IBM Journal of Research and Development* 3 (3): 535–554.

Senior Supervisors Group. 2010. “Observations on Developments in Risk Appetite Frameworks and IT Infrastructure.” New York.

Sparks, Evan. 2017. “Regulatory Compliance? Elementary.” *ABA Banking Journal*, May 1.

Stabile, Carol. 2010. “Data Visualization: Using Interactive Analytics to Combat Financial Crime.” *ACAMS Today* 9 (4).

Stein, Kara. 2016. “A Vision for Data at the SEC.” Keynote address at “Big Data in Finance” conference, Ann Arbor, MI, October 27–28.

SWIFT (Society for Worldwide Interbank Financial Telecommunication). n.d.–a. “Answers to Your Questions about the KYC Registry.” La Hulpe, Belgium. Accessed November 1, 2017.

———. n.d.–b. “Data Standards: BIC (Business Identifier Code).” La Hulpe, Belgium. Accessed August 11, 2017.

———. 2015. “SWIFT Response to CPMI Consultative Report on Correspondent Banking.” La Hulpe, Belgium.

———. 2016. “SWIFT’s KYC Registry Crosses 3,000 Member Milestone.” Press release, November 22. La Hulpe, Belgium.

———. 2017a. “SWIFT Aligns KYC Registry with Updated Wolfsberg Due Diligence Questionnaire (DDQ) for Correspondent Banks.” Press release, October 16. Toronto.

SWIFT and Deloitte. 2012. “Growth, Risk and Compliance: The Case for a Strategic Approach to Managing Reference Data.” White paper. La Hulpe, Belgium: SWIFT.

Thomson Reuters. 2016. “Thomson Reuters Launches Know Your Customer Solution for Africa in Partnership with Barclays Bank, Rand Merchant Bank, and Standard Bank of South Africa.” Press release, July 13. Johannesburg, New York, and London.

———. 2017. “Standard Chartered Bank Announces Intention to Join Thomson Reuters KYC Managed Service for Africa.” Press release, May 15. Johannesburg, New York, and London.

Todd, Sarah, and Marc Hochstein. 2014. “The Race to Build a Know-Your-Customer Registry.” *American Banker*, December 18.

Twiggs, Darryl. 2015. “The Shared Service Utility Model: A Lifeline for the Financial Industry?” *FX-MM* 21 (3): 40–41.

United Nations. 2015. “Sustainable Development Goal 16.”

UN ITU (United Nations International Telecommunications Union). 2016. “Review of National Identity Programs.” Focus Group Technical Report. Geneva.

van Liebergen, Bart. 2017. “Machine Learning: A Revolution in Risk Management and Compliance?” *Capco Institute Journal of Financial Transformation* 45:60–67.

Varian, Hal. 2014. “Big Data: New Tricks for Econometrics.” *Journal of Economic Perspectives* 28 (2): 3–28.

WEF (World Economic Forum). 2016. “A Blueprint for Digital Identity.” Geneva.

Windsor, Richard. 2017. “Intelligent by Design.” *Financial World*, June/July, 7–8.

Wolf, Stephan. 2017. “The True Nature of Identity in Business Applications: Working Paper for the ISO 17442 Review—Embedding the LEI within the Standards Landscape.” Basel, Switzerland: Global Legal Entity Identifier Foundation.

Wolfsberg Group. 2014a. “Wolfsberg Anti–Money Laundering Principles for Correspondent Banking.” Geneva.

——. 2014b. “Wolfsberg Frequently Asked Questions (‘FAQs’) on Correspondent Banking.” Geneva.

——. 2017. “The Wolfsberg Correspondent Banking Due Diligence Questionnaire 2017.” Geneva.

World Bank. 2015a. “Report on the G20 Survey on De-risking Activities in the Remittance Market.” Washington, DC.

——. 2015b. “Withdrawal from Correspondent Banking: Where, Why, and What to Do About It.” Washington, DC.

——. 2017a. “Global Financial Development Report 2017/2018: Bankers without Borders.” Washington, DC.

——. 2017b. “Migration and Remittances: Recent Developments and Outlook.” Migration and Development Brief 27. Washington, DC.

Zarate, Juan, and Chip Poncy. 2016. “Designing a New AML System.” *Banking Perspectives* 4 (3): 26–36.