Fintech and Regtech: Data as the New Regulatory Honeypot

Leon Perlman[1]

Abstract[2]

The emergence of the Internet in the 1990s and the omnipresence of mobile phones in the 2000s as a pivot for everyday interactions between people and businesses, has led a Cambrian explosion in innovation and resultant business models and technologies. Mountains of data are being generated, ready to be used in still further sets of innovations to the degree that 'data' now plays a consequential and pivotal role in the development of innovations in financial ecosystem through as financial technologies (fintech) created by eponymous progenitor financial technology companies ('fintechs'). Fintech has enabled a multitude of innovations, from lending using alternative credit scores, to digital financial services (DFS), to wealth management.

It has also facilitated the emergence of regulatory technology ('regtech') solutions being implemented by regulators and business to ease and ensure compliance as well as act as an early warning for the entities and supervisors alike to events of a financial integrity and systemic nature, such as liquidity crunches and attempts at wholesale money laundering.

The crucible that both fintech and regtech pivot on is this growing mountain of 'big data' and new artificial intelligence algorithms augmented by self-learning 'machine-learning' systems that are able to process and analyze more data at greater speed, accuracy and efficiencies.

Regtech represents a confluence of these activities, where this data is used by regulators for supervisory – 'suptech' – purposes and by supervised entities for their own internal compliance needs in an effective, cost efficient manner.

Both are still at their genesis stage though, with regulators still grappling how to enable fintechs as separate entities, and how available customer data can be used to foster competition and to implement regtech solutions. Similarly, use of artificial intelligence technology to analyze data and undertake predictive analysis on, for example, risk analysis in credit decisions may inadvertently introduce bias in decision making.

This study pieces these disparate – fintech, banking, big data - strands together to identify and analyze regulatory models available for catalyzing fintech, fintechs and regtech, including the potential need for ancillary regulation that would be a touch-point of both regtech and fintech ecosystems to close any potential regulatory gaps and to ensure regulatory certainty in the use of technologies and the surfeit of data powering both fintechs and regtech.

This includes the sourcing use of personal data, cloud computing and data localization/safe harbor rules, sharing of data for anti-money laundering purposes, rules around recognizing data stored on a distributed ledger technology/blockchain as being recognized for evidentiary and other purposes, and use of artificial intelligence and machine learning to analyses in a manner that does not create or perpetuate algorithmic biases and unintended red-lining of classes of people for access to financial services and products.

ABBREVIATIONS

AI	Artificial Intelligence
IoT	Internet of Things.
AML	Anti-Money Laundering
AMLU	Anti-Money Laundering Unit
APC	Asia Pacific Regional Intelligence and Analysis Center
API	Application Programming Interface
BIS	Bank for International Settlements
BSP	Bangko Sentral ng Pilipinas
CBN	Central Bank of Nigeria
CDD	Customer Due Diligence
CFT	Combating the Financing of Terrorism
CFTC	United States Commodity Futures Trading Commission
CIV	Customer Identification and Verification
CNBV	Comision Nacional Bancarias y de Valores
DFS	Digital Financial Services
DFSP	Digital Financial Service Provider
DLT	Distributed Ledger Technology/Technologies
EU	European Union
FCA	Financial Conduct Authority
FI	Financial Institution
Fintech	Financial Technology
FIR	Fair Information Practices
FIU	Financial Intelligence Unit
FSB	Financial Stability Board
FSP	Financial Service Provider
GDPR	General Data Protection Regulation
GFC	Global Financial Crisis
ID	Identification Document
IFC	International Finance Corporation
IP	Internet Protocol
IT	Information Technology
KYC	Know Your Customer
MAS	Monetary Authority of Singapore
MIS	Management Information Systems
ML	Money Laundering
MNO	Mobile Network Operator
MoU	Memorandum of Understanding
OECD	Organisation for Economic Co-operation and Development
PSP	Payments Service Provider
R2A	Regtech for Regulators
Regtech	Regulatory Technology
RFI	Request for Information
RTS	Regulatory Technical Standards
RFP	Request for Proposal
SIM	Subscriber Identity Module
SLA	Service Level Agreement
SME	Small Medium Enterprise
Suptech	Supervisory Technology
TPP	Third Party Provider
TSP	Technology Service Provider
UIDAI	Unique Identification Authority of India
UK	United Kingdom
UNCDF	United Nations Capital Development Fund
USAID	United States Agency for International Development

Contents

1 Introduction

The emergence of the Internet in the 1990s and the omnipresence of mobile phones in the 2000s as a pivot for everyday interactions between people and businesses has led a Cambrian explosion in innovation and resultant business models and technologies. Mountains of data are being generated, ready to be used in still further sets of innovations to the degree that 'data' now plays a consequential and pivotal role in the development of innovations in financial ecosystem through as financial technologies (fintech) created by eponymous progenitor financial technology companies ('fintechs'). Fintech has enabled a multitude of innovations, from lending using alternative credit scores, to digital financial services (DFS), to wealth management.

It has also facilitated the emergence of regulatory technology ('regtech') solutions being implemented by regulators and business to ease and ensure compliance as well as act as an early warning for the entities and supervisors alike to events of financial integrity and systemic nature, such as liquidity crunches and attempts at wholesale money laundering.

The crucible that both fintech and regtech pivot on is this growing mountain of 'big data' and new artificial intelligence algorithms augmented by self-learning 'machine-learning' systems that are able to process and analyze more data at greater speed.[3]

The regulatory dynamic surrounding the emergence of the Internet, the gig 'sharing economy,' and even the emergence of the crypto-economy powered by new forms of crypto-currencies and tokens, has been characterized to a large degree by a cat 'n mouse game – often referred to as 'regulatory dialectics' - between innovators and regulators whereby regulatory action is met by a private sector response designed to ameliorate the impact of that regulation. In some cases, this response may aim to side-step regulations, which may prompt the authorities to tighten the regime further. With the rapid pace of technology innovation, a more apt metaphor may be the tortoise and the hare, characterized by what we term a regulatory-innovation dissonance where regulators and policy makers struggle to keep pace with identifying and understanding new technologies and their regulatory and social impacts.

This dissonance manifests in terms of balancing innovation with financial integrity in so far as allowing fintechs to operate without having to fit within legacy institutional frameworks, nor within the confines of inflexible rules unable to adapt to rapid pace of technology-driven innovation. These institutional and rule-based paradigms may be 'replaced' with a functional approach to regulation where entities are instead regulated on the basis of the services they offer, whilst a more flexible principles-based regime replaced the strict rules-based regimes to more easily adapt to new service offerings.

Often, and now more frequently, these services offerings pivot around aggregation and use of data for any number of purposes. For example, fintechs could use distributed ledger technologies to aggregate and store data from any number of sources and use artificial intelligence to analyze the data to undertake predictive analysis for credit scoring. Data sources for fintechs for these purposes could be via so-called open banking regimes, where various regulatory regimes or market actions allow them secure access to erstwhile closed and proprietorial data held by banks. Or the data source could be as obtuse as call data records held by mobile network operators, allowing fintech to use artificial intelligence (AI) to provide alternative credit scores.

Regtech represents a confluence of these activities, where this data is used by regulators for supervisory – 'suptech' – purposes and by supervised entities for their own internal compliance needs in an effective, cost efficient manner.[4]

Both are still at their genesis stage though, with regulators still grappling how to enable fintechs as separate entities, and how available customer data can be used to foster competition and to implement regtech solutions. Similarly, use of artificial intelligence technology to analyze data and undertake predicative analysis on for example, risk analysis in credit decisions may inadvertently introduce bias in decision making.

This study pieces these disparate – fintech, banking, big data - strands together to identify and analyze regulatory models available for catalyzing fintech, fintechs and regtech, including the potential need for ancillary regulation

that would be a touch-point of both regtech and fintech ecosystems to close any potential regulatory gaps and to ensure regulatory certainly in the use of technologies and the surfeit of data powering both fintechs and regtech.[5]

This includes the sourcing use of personal data, cloud computing and data localization/safe harbor rules, sharing of data for anti-money laundering purposes, rules around recognizing data stored on distributed ledger technology/blockchain as being recognized for evidentiary and other purposes, and use of artificial intelligence and machine learning to analyses in a manner that does not create or perpetuate algorithmic biases and unintended red-lining of classes of people for access to financial services and products.

The contours of an appropriate regulatory strategy are outlined below.

2. 'Fintech' and 'Fintechs'

2.1 Overview

The term "fintech" is a contraction of the words "finance" and "technology" and refers to the technological start-ups that are emerging to challenge traditional banking and financial players and covers an array of services, from crowd funding platforms and mobile payment solutions to online portfolio management tools and international money transfers.

Fintech is a broad term that requires definition and currently regulators are working on bringing out a common definition. The Bank of International Settlements' (BIS) Financial Stability Board (FSB) defines it as:

> '[T]echnologically enabled financial innovation that could result in new business models, applications, processes, or products with an associated material effect on financial markets and institutions and the provision of financial services.'[6]

This definition aims at encompassing the wide variety of innovations in financial services enabled by technologies, regardless of the type, size and regulatory status of the innovative entity. The broadness of the FSB definition is useful when assessing and anticipating the rapid development of the financial system and financial institutions, and the associated risks and opportunities.[7]

Innovations here can relate to economist Josephs Schumpeter's description[8] of innovation as the 'commercially successful application of an idea,' versus invention, the initial development of a new idea, and then from diffusion, the widespread adoption of the innovation. Incremental innovation may create new regulatory touch points.

Although 'fintech' is an umbrella term, we bifurcate it in this study as 'fintech' being the technology catalyst and enabler (functional and ancillary activity), and 'fintechs' as a set of actors (institutional). This bifurcation assists in our later proposal on how to, if needed, regulate 'fintechs' and the data they generate through ancillary regulations that are also coincident with regulations that may be needed to fill regulatory gaps to enable regtech solutions.[9]

Licensed financial institutions and others clearly use not only 'fintech' products developed by external fintech companies but may also develop their own 'fintech' solutions. Ultimately though, consumer choice and market efficiency can be catalyzed by the introduction of new sets of *classes* of product types using fintech that service a particular market function, be that lending, remittance, or investment-related.

Fintech innovations have the potential to deliver a range of benefits, in particular efficiency improvements and cost reductions. Technological developments are also fundamentally changing the way people access financial services and increasing financial inclusion. There is large investment in fintech sector by venture capital funds, with the value of fintech deals worldwide during the first half of 2019 at USD 22 billion, and with the number of global deals increased by two percent to 1,561 compared to 2018.[10]

2.2 Fintech Activities

Some of the major fintech products and services currently used in the market place are Peer to Peer (P2P) lending platforms, crowd funding, distributed ledger (blockchain) technology (DLT), big data generators and analytics, smart contracts, and robo-advisors. These fintech products are currently used in international finance, which bring together the lenders and borrowers, seekers and providers of information, with or without a nodal intermediation agency.

This flurry of activities raises questions over what kind of financial landscape will emerge in the wake of the digital transformation and importantly, how to regulate it.[11] **Exhibit 1 c**ategorizes fintech innovations in relation to financial services. Storing of user funds by fintechs would attract the most regulatory scrutiny, because of safety and soundness and anti-money laundering concerns.

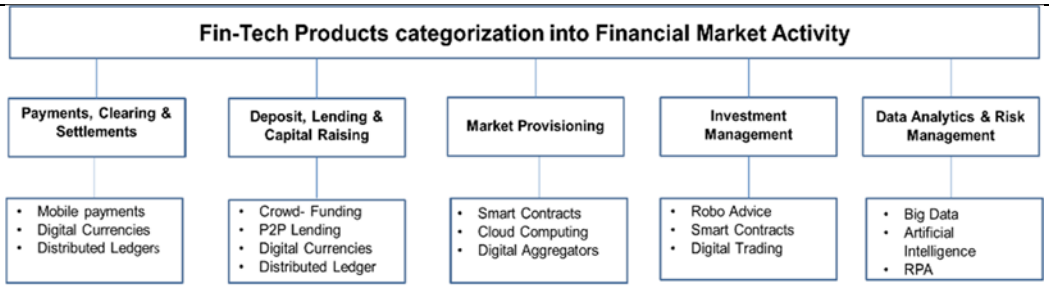


Exhibit 1: Categorization of fintech innovations in relation to financial services. This flurry of innovation from the fintech sector raises questions on how to regulate it. Storing of user funds by fintechs would attract the most regulatory scrutiny, because of safety and soundness and anti-money laundering concerns.

3 Sources and Uses of Data

3.1 Overview

> *"Data is the new oil. It's valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc. to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value."*
>
> - **Clive Humby**, UK Mathematician and apparent originator of the term.[12]

Vast amounts of data are being collected on individuals and machines. Internet use and mobile phones in particular are rich sources of data, provided by individuals freely in exchange for use of nominally free services such as messaging, maps guides, wellness guides, and email.

Data is also generated by connected cars, industrial machines, artificial intelligence, toys and other devices under the rubric of the Internet of Things (IoT). Often this is 'personal data', seen as any information relating to an individual, whether it relates to their private, professional, or public life.

While 'data' and 'information' are used interchangeably in various legal contexts,[13] Data scientists may use the term "data" to refer to discrete, objective facts or observations that are unorganized, unprocessed and without any specific meaning, versus 'information' to refer to data that has been shaped into forms that are meaningful and useful to human beings.[14] 'Big data' and associated 'big analytics' refers to novel ways in which organizations, including government and businesses, combine diverse digital data sets and then use statistics and other data mining techniques to extract from them both hidden information and surprising correlations.

'Data' or connected information can be classed under the rubric of 'big data' and mining this data as big data is the fuel of the digital economy.[15] Stakeholders in digital markets often frame claims, negotiations and controversies

regarding data access as one of ownership. Businesses for example regularly assert and demand that they own data, while individual 'data subjects' also assume that they own data.[16]

This tension highlights serious privacy concerns[17] as de facto, with a few exceptions, companies, not consumers, control the market in personal data with their own interests in mind. [18] More companies are adopting data- driven business models and strategies to obtain and sustain a competitive 'data-advantage' over rivals. [19] Data-driven mergers, like Facebook's acquisition of WhatsApp, have the potential to lessen non- price competition in terms of the array of privacy protections offered to consumers. [20]

Exhibit 2 outlines big data uses in the mobile phone centric DFS[21] in the developing world, particularly through seemingly unrelated and unimportant 'exhaust data' emanating from mobile phone use.

As DFS evolves from its genesis as primarily a remittance-type service to a more transactional offering that includes services such as insurance, investments and credit provision, service providers may want better data sets to assist them to develop new products, to assess customer risk, and to target the correct market segments.[22] For provision of credit, be it short-term micro-credit or a longer term macro-credit product, providers need specific data sets to assess risk and credit worthiness.[23] The data is limited though: only 10% of people in eight sub-Saharan countries, for example, have verifiable online financial data.[24]

For many DFS markets, the most cogent data sets are often those that can be gleaned from mobile phone use, either from conventional telecommunications activity use, through transactional data in DFS or similar transactions obtained by DFS providers such as mobile network operators (MNOs) or through third party smartphone app providers.[25]

In the telecommunications (use) context for example, Call Data Records (CDRs) captured in the course of their operations by MNOs are evolving from simply being flat records of telecommunications service use by individual customers to being the cradle of rich data insights made possible by the connective tissue of big data algorithms. This so-called 'exhaust' data scrapped from these data sources can reveal a lot more on customer behavior, and thus credit worthiness.[26] These metrics are the maximum types of data sets that can be derived from customers with feature phones,[27] augmented however, if the MNO also provides DFS products.

Even richer data sets can be gleaned from users with smartphones, who may use apps that reveal further information about them. For example, some new smartphone apps from DFS credit providers will request and obtain from the user consent to mine their contact lists, get device details, obtain biographical data in registration forms beyond that can be obtained in (often mandatory) SIM card registration, as well as track their calls, text messages such as SMS, instant messages, digital purchase habits, and location.[28] Similar data and results can be obtained by messaging and social network apps who have payment components added, such as those from Tencent's 'WeChat Pay' application in China, and social network behemoth Facebook's 'Messenger' application.

This accumulated data becomes valuable in creating alternate credit scores and in then facilitating provision of credit to some of these profiled users. In many cases, however, users may not be aware data is being scraped and used as a basis for developing an alternate credit score or affecting current credit bureau scoring data. These privacy concerns have garnered the attention of some regulators.[29]

Entities who may be in a position to accumulate data used to create alternative credit scores may potentially use the data to their own advantage by not providing the complete data sets as required to credit bureaus, and/or selectively providing the data only to preferred parties. Entities with significant market power (SMP) may be able to utilize their internal data to the potential detriment of smaller players.

Exhibit 2: Big Data Uses in Digital financial Services

Companies also use what is termed 'screen scraping' to connect the dots on individual. Here, automated scripts collect displayed data elements from one application so that the data can be used by another application. It is often used to collect analytical data from financial transactions, which may require customer-supplied credentials to log in and access the data as if the screen scraper was the customer. The result is that the customer obtains free services, for example, detailed analysis of their health or spending habits. Despite the usefulness of this technology, shortcomings remain in security and speed. Large pools of data can take screen scraping tools 5–10 minutes to retrieve. Passwords and additional security information, once passed to a third party, becomes more vulnerable to loss.[30] While there are no known hacks related to screen scraping, the risks for fraud are mounting though.[31]

Often the data can be access through what are known as Application Programing Interfaces (APIs) - a set of rules and specifications for software programs to communicate with each other to form an interface between different programs to facilitate their interaction.[32] There are also 'reverse engineering' processes to acquire data, where companies analyze applications to extract information about its source code so as to understand the code in order to determine which information is exchanged between an application and a server.[33]

3.2 Data Analytics Using Artificial Intelligence and Machine Learning

While AI and Machine Learning (ML) are seen as one of the most innovative forms of fintech, at the algorithmic level, AI has been around since the 1980's and development has been relatively minor.[34] As AI and ML have been around for decades, the new moniker is simply using more (big) data to model and train algorithms.

Where it has found a renaissance is in the amount of data available for analyzing, coincident with exponential increases in storage and computer processing power, both significantly improved with the emergence of cloud computing power and 'data lake' storage offered by cloud computing providers. That is, AI/ML is now more practical to operationalize and use. The amount of 'big data' sets for use in training AI systems has thus improved, leading to massive use AI/ML cases in everything from finance to health care. Still, newer AI algorithms and ML mean that where AI used to need lots of data for ingestion to analysis, new 'transfer data' and 'fine tuning' techniques reduce the amount of data needed for AI.

Many banks have hundreds of models they use internally, such that a coordination issue presents itself. For example, there is a usually a team that develops the AI software and checkers who validate the models. Anything with advanced analytics may be scrutinized more than say a linear regression.[35] Operational risk relates to models development and validation, mostly the non-mathematical components such as AML and fraud.[36] Earlier models were rule-based, but now use clustering algorithms rather than deep learning using neural networks to reduce false positives and to avoid instances where trends are missed because they do not somehow fit the rule sets.

Many have called for algorithmic accountability: laws governing decision-making by complex algorithms or AI. Algorithms can be used to make, or to greatly affect, decisions about credit, employment, education, and more. Algorithmic decision-making can be opaque, complex, and subject to error, bias, discrimination, in addition to implicating dignitary concerns.[37]

In Australia and Singapore, for example, banks are adopting open banking to make data available for consumers on credit/debit card, deposits, and transaction accounts, mortgage accounts of consumers, and recommended products.[38]

3.3 Open Banking

Traditional banking is evolving into open banking. Open banking regimes - also known as 'open bank data'[39] - are being developed around the world to give fintechs and so-called challenger banks access to customer and associated data usually held in a proprietary and exclusive manner by large financial institutions such as banks. The European Union's Second Payment Services Directive (PSD2) and the Open Banking initiative in the UK and other countries,

for example, try to balance innovation with a competitive pressure on banks to innovate. The rationale is to provide a level playing field for all participants, but at the same time foster an innovative, secure and competitive financial market.[40]

Here, data held by closely banks must be shared in a standardized format with non-bank fintechs or challenger banks once explicit approval of the account holder is given. It also forces banks to provide exact information about every product they offer in a computer-readable format so that interest rates and overdraft fees can be processed by third-party apps and price comparison websites.[41] The data sharing interface between banks and the third parties can be in a standardized Application Programming Interface (API), through some form of 'screen scraping,' or through 'reverse engineering. APIs provide advantages for third parties and customers, including potential improvements to efficiency, data standardization, customer privacy, and data protections. However, some challenges associated with the universal use of APIs remain, including the time and cost to build and maintain APIs when done on a bilateral basis with multiple organizations and the lack of commonly accepted API standards.[42] Screen scraping and reverse engineering are also allowed in most jurisdictions as a method for accessing data where APIs are not available.[43]

In all, sharing of customer-permissioned data by banks with third parties is leveraged to build applications and services that provide faster and easier payments, greater financial transparency options for account holders, new and improved account services, and marketing and cross-selling opportunities.[44] It allows customers to potentially obtain services at better terms, for example, by giving a prospective credit provider or an investor one-off access to 12 months income and spending history to allow the account holder to obtain credit with the third party entity at more favorable terms than with their bank. Anyone using an Open Banking service will not need to share their banking login or password with anyone but the bank.[45] Direct customer application allows them to see accounts at various institutions at the same time in a dashboard without having to do multiple logins at different bank web sites. For e-commerce use, it also makes it possible for customers to pay directly from a bank account rather than through a third-party intermediary, or even Visa or MasterCard. The initial data-sharing scope only includes current account data, with credit cards and other payment accounts added later. In the UK regime, only startups that have been approved by the Financial Services Authority (FSA) regulator will be allowed to use the system.[46]

A comprehensive open banking framework can include rules, standards and/or industry practices across a range of issues, as well as different regulatory authorities, especially where unregulated third and fourth parties gain access to bank customer-permissioned data. Regulatory approaches to Open Banking are discussed further below.

3.4 Data Protection Schemes

Big Data and Big Analytics also raise many legal, moral, and ethical issues, such as cyber-security and the accountability of firms for use of algorithms in automated decision-making.[47] Where vast amounts of personal data are shared and transferred around the globe instantaneously, it becomes increasingly difficult for people to maintain control of their personal information. This is the role of 'data protection' as the consort of practices, safeguards, and binding rules put in place to protect personal information. It also ensures that individuals remain in control of it by being able to decide whether or not they want to share some information, who has access to it, for how long, for what reason, and be able to modify some of this information.[48]

By the end of 2019, almost 100 jurisdictions had some sort of data protection laws, built primarily around the Fair Information Practices ('FIPs') which establish a number of individual rights, including access, disclosure, and correction rights, along with general obligations respecting data gathering, storage.[49] This concept originated in the United States, but has become the international standard for data protection, under the Organisation for Economic Co-operation and Development (OECD).[50] The EU's General Data Protection Regulation's (GDPR) primary principle is that consumers own and control their data, while other jurisdictions' data privacy laws are premised on the principle that consumers, including banks, 'own' the data they maintain. The GDPR establishes a system of generally applicable notification and access rights,[51] starting from collection of personal information from an individual, wherefrom a company must provide the purpose for which data is gathered, the recipients of the data, and the retention period of the data. Nearly identical information must be disclosed if a company obtains personal data not directly from an individual but from another party.[52]

Often though, there gaps in data protection laws are exposed when faced with technological innovations or contradictions with, especially, financial sector laws which are discussed below. These may or may not be cured by ancillary laws and regulations that are sector-specific or principles-based. In their absence, two motifs may inform schemes for data protection: Data protection by design and data protection by default.[53] The former employs pseudonymisation to replace personally identifiable material with artificial identifiers, as well as encryption to encoding data so that only those authorized can have access to read the data. Companies/organizations are encouraged to implement technical and organizational measures, at the earliest stages of the design of the processing operations, in such a way that safeguards privacy and data protection principles right from the start. In the latter, companies/organizations should, by default, ensure that personal data is processed with the highest privacy protection so that by default personal data isn't made accessible to an indefinite number of persons. This may curtail data for processing, storage and accessibility. In a social media context for example, users' profile settings should, by default, be set in the most privacy-friendly manner. [54]

3.5 Cloud Computing and Data Localization

Many fintechs want to use cloud computing providers simply as an extension of their own servers, storing data in an encrypted form that does not expose customer data to the cloud provider. The types of cloud services available include:

- IaaS - Infrastructure as a Service
- PaaS - Platform as a Service
- SaaS - Software as a Service
- BaaS - Backend as a Service

Because data is in the cloud, there can be no single point of failure, which is nirvana for especially developing countries given many instances where there is a lack of reliable power or natural disasters. Even where there are regulations allowing cloud computing of some form, fintechs see any potential requirement to keep data in a local cloud provider as impractical and would rather have the option to use major international cloud providers such as Amazon, Microsoft or IBM.

3.6 Storage and Distribution of Data Using Distributed Ledger Technologies

The emergence of Distributed Ledger Technologies[55] such as blockchain has unleashed an entirely new data ecosystem predicated on a decentralized mechanism of control and access, with (ostensibly) no single entity controlling a DLT.[56] That is, for any of its protocols, instead of being 'dumb' pipes that simply carry data and valuable applications above them, with DLTs the value can be and is embedded *inside* the protocol itself. One could now call this shift the equivalent of the 'internet of value,' or Web 3.0 as some term it.[57]

In this latest transformation, centralization is replaced by protocols that facilitate and allow data - and thus innate, embedded value - to be distributed. In a nod to the decentralization motif of DLT, this could be without a central control point mediating what can be sent, used and seen.[58] Disassembly of the components of the DLT protocol demonstrates that there are two components at play: the technical parts that mediate the interactions with other users of the protocol (the nodes) and the business end called 'tokens' that - depending on the DLT protocol - are entirely programmable, even as a form of 'programmable money.'

4 Regtech

4.1 Overview

Extensive regulatory reforms imposed as the result of the Global Financial Crisis (GFC) of 2007-2009 have caused dramatic structural changes in finance around the world. The GFC led to an internationally coordinated process of regulatory reform, focused on reducing risk-taking and systemic risks in the financial sector.

These reforms have also been a major driving factor in the adoption and use of new technologies in the sector, particularly the technologies that aid compliance with regulation – a concept known as 'regtech.' In parallel with, and increasingly coupled to, these financial regulatory reforms have occurred in extensive reforms of data protection, the advent of open banking, and the development of digital identification regimes.

Together, these factors form a regulatory ecosystem that supports a transformative transition from traditional banking and finance to data-driven banking and finance. From the 1960s to the cusp of the global financial crises of 2008-2010,[59] there has been a large growth in financial institutions in both size and scope. The complexity of operations for financial institutions and product mixes increased,[60] becoming more quantitative and technology-driven,[61] driving the emergence of complex regulations[62] and similarly increasing associated compliance costs.[63]

Improvements in computer processing power and improved software solutions have allowed financial institutions to adapt to the increasing burden of regulatory requirements, however, concomitant technology to supervise and facilitate compliance has arguably not evolved in any significant way since the 1990's.[64] That is, baseline compliance and supervisory reporting tools[65] are still largely Excel,[66] XML[67] and email-based for submission of data to supervisors.

Analysis of collected data by supervisors is also largely manual with little feedback available to check whether the requisite data has been provided, whether the data is in the correct format, whether it is accurate, and whether any specific follow up or supervisory actions are needed.

Also, the reporting paradigm is the same: the supervised entity fills in a spreadsheet, sends it to the supervisory authority, who then checks the data and sends any queries to the financial institution. The process then repeats. Data analysis is usually a separate process with its own variances.

With larger numbers of entities and products to monitor and supervise, the data and supervisory burden on supervisors has exploded. Use of more automated and innovative technology solutions for compliance and supervision has emerged in the concept of specific technology solutions, or regtech, first defined by the United Kingdom's Financial Conduct Authority (FCA) as:

> *"[A] sub-set of fintech that focuses on technologies that may facilitate the delivery of regulatory requirements more efficiently and effectively than existing capabilities."*[68]

As fintech/s and techfin/s[69] disrupt the financial industry,[70] regtech is being fueled by the rapid technological developments and disruptive innovation in fintech.[71] The underlying technologies of fintech[72] were used and are being used in a regulatory context to drive regtech innovation evolving from the need to reduce compliance[73] costs for financial institutions, to being adopted by financial service providers (FSPs),[74] fintech companies, central banks and also businesses in other industries for other purposes.[75]

Central banks in particular are exploring new ways to use these new technologies for onsite and real-time analysis. Adoption, however, requires a team with technical, policy making and supervision expertise to spearhead the initiative.

4.2 Regtech for Financial Market Participants

Supranational Approaches: Governments and international bodies brought about major regulatory changes that increased capital requirements and compliance costs for financial institutions after the global financial crisis (GFC) of 2008.[76] The G20 in 2009 established a financial standard setting body (SSB), the Financial Stability Board, to play a key role in promoting the reform of international financial regulation and monitoring of the international financial system for any signs of systemic weakness.[77]

Financial Institutions: Financial institutions around the globe struggled to different degrees post-GFC with increased compliance burdens and monitoring and restrictions on investments, which catalyzed the need for development and adoption of regtech as means to reduce the cost of compliance and to manage risk.[78] Uncoordinated timelines and agendas for the implementation of overlapping regulations and constant evolution of regulation[79] furthered costs and complexities.[80] This, coupled with the lack of trust in the financial system led banks in developed countries to become more hesitant to providing credit and maintaining relationships that provided low returns and higher risks. Many developing countries were affected by this credit-freeze and low risk appetite.[81] This affected trade, remittance flows, aid and capital inflows in developing countries.[82]

DFSPs: While increased regulation for banks limited their scope but allowed less regulated non-bank digital financial services providers (DFSPs) to grow.[83] In developing countries, growth of DFS targeted – and still targets – financial inclusion and economic development.[84] Rapid pace of DFS innovation and introduction of new customers and providers in the market, however, gives rise to newer risks. Risks related to data privacy and consumer protection can be more pronounced in developing countries due to low financial literacy, lack of appropriate policies and regulations, underdeveloped technology ecosystem and weak infrastructure.[85]

Compliance burdens are especially heightened in the provision of DFS where multiple regulators including financial and telecommunication regulators are involved, leading to duplication in DFS reporting requirements within the same authority and for multiple authorities.[86] DFSPs may need to invest time, skill and money into compliance activities which can be difficult if they have limited resources. Such compliance burden could force DFSPs to compromise on innovation.

A common response to high compliance burdens has been to increase the size of financial institution's risk management and compliance teams.[87] While this may be a solution for some, it may not be feasible for smaller DFSPs – usually start-ups – with limited financial and human resources. DFSPs can adopt cost cutting regtech solutions, either developed in-house or by technical service providers (TSPs) that tackle different aspects of regulatory issues.[88] These include issues related to market and staff surveillance, reporting, anti-money laundering/combating the financing of terrorism (AML/CFT),[89,] know your customer (KYC),[90] customer due diligence (CDD)[91], and risk management.[92]

The gap between DFSPs and regulatory requirements may be bridged by regtech and in the process, it can increase access to underserved populations.[93]

4.3 Suptech for Financial Regulators

Unlike legacy technologies[94] and associated manual processes that have been used by regulators for their own internal assessments and supervisory remits, regtech can facilitate the collection and organizing of high velocity, diverse types and large volumes of data in agile, fast and integrated ways to facilitate automated extraction of actionable data.[95]

A key attribute of regtech is the 'check' function, which acts as a feedback loop to determine whether reports have been submitted on time, accurately, in the correct format and to the correct supervisor.

Supervisory functions: Regulators may require financial and operational data from market participants to produce statistics that drive their understanding of the market and policy decisions. Some types of data collected by central banks and other DFS supervisors include:[96]

- Financial statements (balance sheet, cash flow, income statement)
- Financial ratios (liquidity ratios, capital adequacy ratios)
- Volume and value of transactions
- Number of transaction points
- Number of accounts and total balances

- Description of frauds and actions taken, actions taken on consumer complaints, risk management practices and IT systems
- Losses from frauds, consumer compensations

With the adoption of regtech by market participants, they may be able to report data more frequently, monthly, daily or even real-time, making large amounts of data accessible to regulators. Regulators may then be able to use regtech to process and analyze the data.

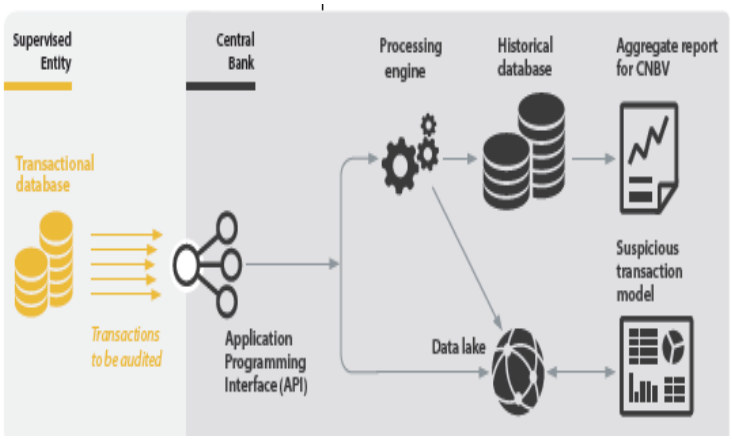
Internal Processes: For central banks specifically, regtech can potentially improve efficiency and effectiveness of their internal processes as well as external processes involving both supervised entities and the central bank.

Regtech has assisted central banks to address the challenges of monitoring a rapidly evolving financial sector that lacks proper tools and infrastructure for supervision and monitoring[97] by providing alternative processes. It may also allow central banks to develop appropriate regulations, by facilitating better understanding of new market participants and technologies.

Many central banks in developing countries, however, face unique challenges that may hinder the adoption of regtech. Without regtech, central banks may not have the capacity to monitor the new additions to the financial system, so they are more likely to impose stricter regulations to deal with the new and unknown risks posed by the changing financial landscape. DFSPs could hence face regulatory uncertainties and compliance burdens as central banks try to balance innovation and stability.

A suptech prototype used by the Mexican banking regulator Comisión Nacional Bancaria y de Valores (CNBV)[98] represented a fundamental re-engineering of its approach to AML supervision.

The manual process is shown on the top right, the new regtech solution at the bottom right. The new regtech solution consists of a mix of open-source, cost-effective (that is, no licensing requirements), best-in-class technologies targeted at the various pain points of the existing reporting system. What had previously taken days and weeks to collect could now be achieved in mere seconds. Together they formed a coherent, streamlined architecture for the transmission, processing, warehousing, and analysis of banks' transactional data.



Application programming interface (API): The API establishes a secure, direct line of machine-to-machine data transmission between the supervised institutions' transactional databases and CNBV's processing engine. Raw data is "pushed" or "pulled" directly to CNBV's systems, providing the supervisor with the raw data as well as select suspicious transactions, obviating the need for manually-populated spreadsheet templates, insecure email transmissions, or time-consuming CD submissions.

Processing engine: A processing engine receives the data and instantly runs validation tests in order to verify the quality, content, and structure of the reports. Incorrect or incomplete reports are automatically rejected, having a single processing engine ensuring that all tests are run uniformly rather than on separate spreadsheets whose formulas may be inconsistent, broken, or out of date. It also allows for more complex number crunching than might be possible in Excel and significantly cuts down on processing times

Database: The processed data is funneled into a "data lake" for storage together with the original raw data and the AML model output data. Retaining all the information that is streaming through the CNBV's platform ensures that all transformations can be traced back to the original source. The lake also serves as a staging area for the historical data warehouse, the database where the "treated" data is stored for reporting and analysis. The flow of data throughout the platform is controlled automatically through edge computing formulae, which cuts out manual workarounds and further enhances the efficiency of the solution.

Exhibit 3: The Mexican financial regulator CNBV's AML suptech data architecture regtech solution compared to the manual process.[99]

4.4 RegTech and Fintech

The *intersection* between regtech and fintech may arise within the data sourcing, storage (in data lakes) and use and any ancillary regulation that may be needed to address either new technologies as a whole, and/or their impact. For example, use of AI, ML DLTs, AML, cloud computing and data localization, data protection rules relating to proper use of stored and accumulated data, and cyber-security. **Exhibit 4** shows the commonalities in regulation where regtech and fintech intersect through data lakes and regulation, where needed, of data storage and use.

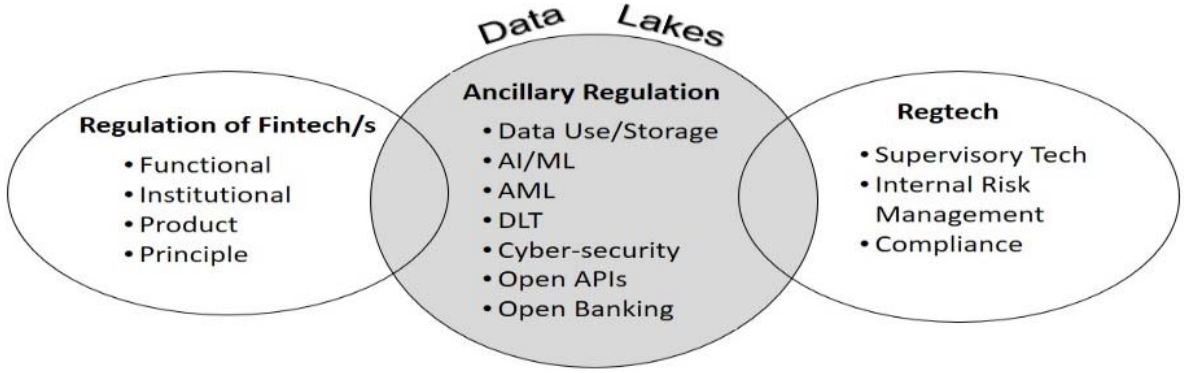


Exhibit 4: Regulation of fintech (and fintechs) and regtech. The graphic shows the commonalities in regulation where regtech and fintech intersect and which may require ancillary regulations to recognize, process, store and use the surfeit of data generated by fintech and regtech. Regulation of fintechs and data is discussed below in Section 5 and Section 6 respectively.

Fintech, by definition, also involves only the financial sector, whereas regtech can apply in any area of regulation, compliance and system design, whether in the context of finance or otherwise.[100] At the same time, rapid evolution in fintech is raising new risks. The sheer amount of data facilitates looking at correlations rather than causations and correlations can lead to unintended and socially regressive consequences. Yet the methods to properly supervise and control self-learning algorithms are yet to be developed.

Regtech differs from fintech in that fintech mostly addresses business processes and technologies, as well as 'fintechs' as a class of entity implementing a fintech solution.[101] These could include startups, small medium enterprises (SMEs), and banks.

With the emergence of large data lakes full of 'big data,' regtech and fintech intersect within these data lakes, and form the foci of new sets of regulatory approaches and regulations. Regulation of fintechs and data is discussed below in Section 5 and Section 6 respectively.

5 Policy and Regulatory Approaches to Fintech and Fintechs

5.1 Overview

Regulatory uncertainty surrounding fintech could potentially hamper its development.[102] Faced with the profound changes that fintech is bringing to the banking and financial sectors, regulators are faced with a conundrum – needing to balance their regulatory approaches between (a) regulations that create such barriers to entry that innovation is stifled and incumbents are unduly protected versus (b) a lighter touch which can enable any number of newcomers hindering consumer protection.[103]

Adding to this conundrum are the range and sized of entities, the mosaic of business models, and the classification of the various types of activities, products and transactions.[104] These activities – also called 'verticals' – may be cross-sector, much like DFS has been for financial services.

That means that one entity with ostensibly one service type may, through the use of various technologies simultaneously, touch on a number of regulatory domains, with any number of regulators and rules fastening on it and its services. An entity doing lending based on using so-called 'exhaust data' derived from a customer's mobile phone use to create alternative credit scores, and using a cloud-based platform housed in a foreign jurisdiction and using access to customer credentials based on an 'open banking' regime, for example, may touch the banking, credit, privacy, consumer protection, telecommunications, and consumer protection regulators and any number of regulations imposed by each of those regulators.

5.2 General Principles and Approaches to Fintech Regulation

While fintechs can play an outsized role in the evolution of a country's financial ecosystem, the breadth of the sector and its offerings can make it difficult to talk about 'fintech regulation' *per se,* and indeed fintech as specific entities that are recognized as a specific class and requiring institution-based regulation.

Simply put, if it's difficult to define it and its scope, it's difficult to regulate it. The challenge becomes even more pronounced with technology leaps – such as DFS, DLTs, crowd-funding platforms and AI – which with their breadth of application, invariably challenges the perimeters of many regulatory remits and scope of laws and regulations.

That is, multiple laws, regulations and regulators may be impacted by the emergence of new technology beyond its obvious initial use – or may not be impacted at all given that the tech is new, reflecting that evolving technology is often just an enabler of a particular activity or function. Regulators diverting their resources, though, to understand every new technological innovation could result in inefficient outcomes for regulators and industry.

From the innovators' perspective, regulation may spur what has been termed 'compliance innovation' or 'circumventive innovation.'[105] The latter occurs when the scope of the regulation is narrow and the resulting innovation allows an innovator to escape the regulatory constraints. This may be, for example, because of a narrow-rules-based approach. The former may occur when the scope of the regulation is broad and the resulting product or process innovations remain within the scope of the regulation. This may be the result of a principle-based approach. There may not, however, be a deliberate strategy by the innovator to embrace one or other of these approaches; new technologies and resultant combinations of products it catalyzes may result in either of the approaches being touched upon.

From the regulator and policy–maker perspective, the main current concerns arise not from the technology itself but from the question of *who* is applying technology to finance along with the *speed* of development. Lack of proper understanding by the regulators may lead to onerous conditions being imposed and/or delays in obtaining approvals. In general, provided that the innovator can demonstrate how it intends to mitigate the risks to the public and the risk of its innovation being used for money laundering or the financing of illicit activities, regulators could allow the innovation to proceed to market.

Providing clarity as to the position of fintech companies in the post-fintech regulation era, in particular, is the manner in which a sector, effectively operating in an unregulated environment, will transition into a regulated environment.[106]

5.3 Regulatory Approaches To Fintech Enablement

A number of regulatory approaches are being used by regulators around the world in general. At one end is the *institutional* approach (*who* is providing services) anchored in predictable, but relatively inflexible *ruled-based* regulation. On the other end there is a product-like approach anchored in flexible, *principle*-based regulation known as the *functional approach* (*what* services are being provided).[107]

Some jurisdictions with one or other of these approaches also include a product-based approach that address specific product types, rather than product segments. Each approach and their relative merits are discussed below.

Rules Based:

- Description: Regulators issue largely inflexible rules as to who can provide services, if at all as well as what serviced can be provided.
- Advantages: Creates certainty as to what a fintech must do to comply since the regulator must forward-engineer the implications of compliance for the intended regulatory outcomes, usually via a regulatory impact assessment. For fintechs, the legal predictability (and thus) higher compliance costs associated with a rule-based model may be balanced with the certainly being more attractive to investors. This approach is more likely to create a barrier to entry for subsequent new competitors to existing fintechs.
- Disadvantages: While there may be clarity about the compliance process, overall regulatory objectives may be ambiguous.[108] Compliance obligations can limit the incentive of the supervised entity to do more because the obligations are perceived as sufficiently comprehensive. Compliance costs can be very high as there may be a one-size-fits all approach that is unsuited to a startup.

Principles Based:

- Description: Regulators provide principles fintechs need to abide by rather than specific rules. May be combined with a tiered-based approach.
- Advantages: There is clarity about the regulatory objectives. Also allows fintech startups to provide services without inflexible operating requirements of the rules-based approach. With the tiered-approach, as the startup grows, it faces higher levels of regulatory scrutiny. As the fintech start-up matures, it grows in its capacity and so does its compliance culture, with increasing access to sufficient financial resources.
- Disadvantages: While there is clarity about regulatory objectives, translating these principles into rules that will not trigger compliance liabilities is ambiguous. The flexibility of a principle-based allocates sufficient discretionary power to the regulator to potentially create a level of uncertainty as to what exactly is expected in terms of compliance. That is, whilst a principle-based approach may provide a start-up with the benefit of flexibility at an early stage, this may create limitations in terms of scalability of a business.

Institutional:

- Description: Only specified entities can provide specified services, a one size fits all approach.
- Advantages: For regulators, they can tailor their regulatory capacity needs according to a set number of entities that they may need to regulate based on finite descriptions of entity types and functions. For larger institutions, it keeps out competition as there is usually a high barrier to entry.
- Disadvantages: Entities that do not fit within the finite number of entity types may not be authorized or licensed. A one-size-fit all required collateral and/or license fee will exclude most startups who do not have the funds for this purpose.

Functional:

- Description: Entities, no matter who they are, that fit a functional description of a service or vertical can provide services.
- Advantages: Allows classes of entities performing certain general activities (and vertical) as specified generally to be authorized or licensed by the regulator (if at all needed), and is usually not based on any particular technology, class of or on the size of entity. Entities that fit within a certain level can simply gain authorization.
- Disadvantages: Requires that the regulator to undertake continuous market studies to determine potential functions that can be included in a regulation. A number of regulators may be impacted by the broad functionality, which will induce regulatory arbitrage if regulators do not coordinate on which regulator has specific oversight.

Product Based:

- Description: Regulation is based on the exact product rather than a class.
- Advantages: Allows classes specific products to be authorized or licensed by the regulator (if at all needed), and is usually not based on any particular technology, class of or on the size of entity. Entities that provide the product type within a certain level can simply gain authorization.
- Disadvantages: Requires that the regulator undertake continuous market studies to determine product portfolios. A number of regulators could be impacted by the product however, which will induce regulatory arbitrage if regulators do not coordinate on which regulator has specific oversight. Requires that there be a specific rule for a specific product type or example.

Regulatory actions may vary from "disclosure" to "light-touch regulation and supervision" to a "full-fledged regulation and supervision", depending on the risk implications. A tiered approach could be used, increasing oversight as an entity grows and its risk profile changes.

The contours of an appropriate regulatory strategy are outlined below.

5.4 Supranational Approaches to Fintech Innovation and Regulation

Recognizing the importance of fintechs in the development of financial ecosystems and their role in fostering innovations and competition, in October 2018 the International Monetary Fund and the World Bank Group launched the Bali Fintech Agenda,[109] a set of 12 policy elements aimed at helping member countries to harness the benefits and opportunities of rapid advances in financial technology that are transforming the provision of banking services, while at the same time managing the inherent risks

.

The Agenda's 12 high-level issues are for countries to consider in their own domestic policy. They cover topics relating broadly to enabling fintech; ensuring financial sector resilience; addressing risks; and promoting international cooperation:

- Embrace the promise of fintech.
- Enable new technologies to enhance financial service provision.
- Reinforce competition and commitment to open, free, and contestable markets.
- Foster fintech to promote financial inclusion and develop financial markets.
- Monitor developments closely to deepen understanding of evolving financial systems.
- Adapt regulatory framework and supervisory practices for orderly development and stability of the financial system.
- Safeguard the integrity of financial systems.
- Modernize legal frameworks to provide an enabling legal landscape.

- Ensure the stability of domestic monetary and financial systems.
- Develop robust financial and data infrastructure to sustain fintech benefits.
- Encourage international cooperation and information-sharing.
- Enhance collective surveillance of the international monetary and financial system.

6 Policy and Regulatory Approaches to Data Use in Fintech and Regtech

6.1 Overview

Following a recalibration of laws, regulations and remits following the emergence of e-commerce in the 1990s and then fintechs in the 2000s, a similar reassessment is underway for the singularity of 'data.' Data of all forms is easier to gather, extract, extrude, store and analyze, but how and whether to manage this are still national and regional policy and regulation.

The sheer sources, volume and uses of data challenge policy makers and regulators in assessing interaction with existing laws and regulations.

A number of dedicated data protection laws are emerging in many jurisdictions, although these vary in scope and regulatory oversight. They often include however fundamental consent and privacy expectations, as well as data security requirements,[110] the most far-reaching one being the GDPR in the EU.

The need though for a dedicated data protection law pivots around public policy considerations, not in the context of the rise of surveillance states where there is an abundance of data – from biometrics, spending habits to social standing - collected on citizens, but primarily private companies that wield enormous power in data collection, aggregation and commercial use. Indeed, while there appears to be a strong divergence in thought of the use of data by governments, there appears to be increasing consensus around placing limits on the use of data by the private entities.

Dedicated data protection laws – particularly those that are rule-based – may, however, not be applicable to all use cases and indeed may be able to interact with existing laws and regulations and emergent technologies. The emergence of data lakes raises fundamental questions though on how data regulation – in whatever form it may be – comports and interacts with financial regulation, where often data regulation may be out of step with financial regulation, creating regulation gaps and arbitrage. This may even lead to consumer harms or threats to financial integrity.

Ancillary regulations touching on data acquisition, distribution, storage, use and analysis may be needed to close any potential regulatory gaps, ensuring regulatory certainly in the use of technologies and the surfeit of data powering both fintechs and regtech. That is, these ancillary regulations, as needed, would address the intersection of fintech, regtech and data regulation in use of critical data sets or lakes that are at the core of fintech and regtech.

6.2 Regulatory Approaches

As noted above, an increasing number of jurisdictions are developing dedicated data protection laws for aggregation, storage and use of customer data. They range from omnibus provisions that give individuals the power to dictate what data can be shared by a 'data controller' such as a bank, with whom it can be shared, and for how long, to simple consents to receive information.

Given that data protection laws may not cover every eventuality, technology or circumstance, and to support a functional, principles-based approach to fintech regulation regulations (as needed) for ancillary-type services would need to be developed, or existing regulations clarified to allow authorized/licensed fintechs to use them in furtherance of their activities and verticals. Data-related components within current financial regulations or laws may need to be updated, clarified or developed to avoid regulatory ambiguity, gaps and arbitrage.[111]

At a more granular level, three types of data regulation – particularly relating to data sourcing, sharing and use – are emerging. As a barometer, regulators have either taken or are considering a range of actions related to open banking in their respective jurisdictions.

Some jurisdictions have taken a prescriptive or regulated approach, requiring, for example, banks to share customer-permissioned data and requiring third parties that want to access such data to register with particular regulatory or supervisory authorities.

There is also a facilitative or supervised approach whereby regulators issue guidance and recommended standards and release open API standards and technical specifications.

Others follow a market-driven approach, with no explicit rules or guidance that require or prohibit the sharing of customer-permissioned data by banks with third parties.[112] There are benefits and challenges with each approach when balancing financial system safety and soundness, encouraging innovation and ensuring consumer protection.

While these approaches may be a response to the 'regulatory dialectics' dynamic identified earlier, some of the more proactive measures are clearly designed to enable data sharing – albeit at a glacial pace – for pro-competition and financial integrity purposes, whilst a at the same time keeping an eye on consumer protection.

Clearly, with complexity arising from innovations, regulators cannot foresee all the permutations and their regulatory impact. That leads back to the dialectical approach, whereby regulators without having the ability to provide a fulsome regulatory response can, at best, use ancillary regulations – if at all needed at a public policy level – to close regulatory gaps that may manifest in the *intersection* between regtech and fintech.

These regulations would address either new technologies as a whole and/or their impact, for example, use of AI, ML, open banking, DLTs, crypto-assets, AML, cloud computing and data localization, data protection rules relating to proper use of stored and accumulated data, and cyber-security.[113]

Some of these ancillary regulations are described below. Often though, regulatory forbearance would be suited to particularly circumstances lest regulators fall into the trap of policy-making through enforcement.

6.3 Regulatory Coordination

Within each jurisdiction, multiple authorities can have a role in addressing issues relating to sharing of customer-permissioned data by banks or fintechs with third parties owing to the multi-disciplinary aspects of fintech and data sciences. Relevant authorities may include, for example, bank supervisors, competition authorities, and consumer protection authorities, among others. These are further outlined in **Exhibit 5.**

Regulator	Potential Data-related Remit
Bank Supervisor	Sets requirements and supervises regulated banks.
Technical Standards Setting Body	Establishes standards and certifies entities that comply with set national or international standards.
Telecommunications Authority	Setting technical standards and use of 'exhaust data 'from mobile phone use described above.
Competition Authority	To monitor, promote and, when necessary, take action to ensure well-functioning markets.
Consumer Protection Authority	Ensure consumers are generally not disadvantaged by monopolistic and oligopolistic practices by organizations. In some jurisdictions, their mandates may include ensuring consumers are not disadvantaged by unfair, deceptive, or abusive acts or practices.
Data Privacy Authority	Sets requirements relating to protection of personal and/or customer data.
Alternative Dispute Mechanism	Provides a platform or process to mediate disputes between consumers and organizations.
Others	Any other body that has a mandate over entities engaged in open banking.

Exhibit 5: Regulators potentially involved in data-related regulation.[114]

6.4 Laws and Regulations Supporting Fintechs and Data Sharing

6.4.1 Overview

The ability to share data within the realm of harmonized or certain regulatory regimes is predicated on coordinated and robust data protection and data storage regimes. Because of the evolving nature of the 'big data' and AI paradigms and its cross-sector application, these ancillary regulations are not neatly placed with a specific framework but are mostly a patchwork of regulations that fasten on fintechs and banks.

Such certain or harmonized regulatory regimes are, however, not prevalent, clear or certain in most developing countries where 'ancillary' regulations buttressing direct financial regulations – such as for payments and banking - are for the most part absent. These would, as noted above, relate *inter alia* to regulations on AML, cloud computing, and AI. A number of sandbox and similar environments are percolating out as resource centers that keep track of fintech developments and assist fintechs in navigating regulatory frameworks.[115] In most OECD countries, some but not all of these issues – such as allowing use of cloud computing in financial services – are settled. Only the EU appears to have a coordinated mechanism where there is cross-referencing of each of the various types of ancillary regulations, if all exist. In all though, it is incumbent on the fintech to determine what direct financial laws and regulations, and the ancillary laws and regulations apply to them. **Exhibit 6** shows potential ancillary regulations relating to sourcing, storing and use of data.

Ancillary Regulation	Scope
Consumer Protection	Data protection laws may be required simply to provide explicit consent (and legal certainty) for service providers and banks to store personal data on cloud-based services if current banking regulations only allow this data to be stored on local bank servers.
Data Localization	'Data localization' rules on whether and what data can be stored or shared in cloud servers in foreign jurisdictions and what such 'safe harbor' provisions may entail.
AML Use	Allowing sharing of customer data for AML purposes without violating current prohibitions on 'tipping off,' that someone is being treated as suspicious in terms of AML laws and regulations.
Use of Distributed Ledger Technologies	Determining **w**hether blockchain (distributed ledger technology) use is certain,[116] alongside the legal certainty of 'smart contract' and evidential use
Cloud Computing	Data storage and sharing in cloud-based systems.
Artificial Intelligence	Rules around AI use, for example preventing bias in decision-making, and disclosure of non-humans use in one-one-one customer interactions.
Sandbox and Fintech Labs	Sandbox and fintech labs, which allow fintechs a limited, but safe harbor testing environment for their innovations. The sandbox approach allows new products to be tested in a specific environment without regulatory burdens.[117] Labs are dedicated to supporting fintechs.[118]

Exhibit 6: Potential ancillary regulations relating to sourcing, storing and use of data, even where there is a dedicated data protection law.

Ancillary regulations, however, are in many cases not necessarily required to allow data sharing, but instead may provide legal and regulatory *certainty* where there is no clarity, or where there is even regulatory arbitrage such that multiple – and possibly – conflicting regulations may apply to a service. In many cases regulatory forbearance would suffice to provide clarity.

6.4.2 Data Protection Laws and Regulations

Consumer pessimism about online privacy may have contributed to the development of what has been termed a 'dysfunctional equilibrium,' whereby the market underprovides privacy protection because consumers do not believe that they have control over privacy or that companies really will protect their privacy.[119]

Here is where the so-called data protection laws step in to regulate where they can the use of 'personal data' by organizations to protect certain rights of individuals – organizations are not free to use personal data at will. These protect information relating to an identified or identifiable natural (living) person and in some circumstances, could include information such as Internet protocol (IP) addresses and communications content.[120]

Some 107 countries - of which 66 were developing or transition economies - have passed legislation to secure the protection of data and privacy. In this area, Asia and Africa show a similar level of adoption, with less than 40% of countries having a law in place.[121] Globally, there is an increasing growth in data protection laws, many of which have been modelled on comprehensive guidelines or regulation such as the EU's GDPR, or the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data.[122] **Exhibit 7** outlines of the scope of the EU's GDPR.

The GDPR is a technology-neutral law that establishes a system of generally applicable notification and access rights.[123] For example, from collecting personal information from an individual, a company must then provide the purpose for which data is gathered, the recipients of the data, and the retention period of the data. Nearly identical information must be disclosed if a company obtains personal data not directly from an individual but from another party. It also contains affirmative access rights for individuals, including access to the source of the data and a copy of the data itself.[124] Access rights may be invoked at intervals and give individuals the ability to regularly check in about what information a company has about them, beyond the moment at which data has originally been obtained. It endorses self-regulatory instruments, such as certification and technical standards[125] and specifies that notices to individuals must be communicated "in a concise, transparent, intelligible and easily accessible form, using clear and plain language…"

Profiling using algorithmic decision-making as a subcategory of data processing triggers additional rights under the GDPR.[126] Here, a company must proactively notify individuals of the existence of solely automated decision-making based on profiling, regardless of whether or not that decision-making is solely automated.[127] This must also explains how profiling works.[128]

Exhibit 7: Outline of the scope of the EU's GDPR

As noted by the BIS,[129] some frameworks view banks, and sometimes third parties, as the data owner, but limit their rights to control the use of such data to the boundaries of the consent provided by the customer. Many jurisdictions' consent rules also place restrictions on downstreaming data to fourth parties and on reselling customer data for purposes beyond the customer's initial consent.

The GDPR harmonizes previous legal frameworks data protection, which was fragmented across Member States.[130] Individuals now have the power to demand companies reveal or delete the personal data they hold and requires businesses to be more accountable to the people whose data they collect. All businesses handling EU citizens' data, whether based in the EU or outside, must comply with GDPR, subject to financial sanctions for non-compliance.

While the EU's General Data Protection Regulation (GDPR) has as its primary principle that consumers own and control their data, other jurisdictions are premised on the principle that entities own the data and that permission is required before data is shared by the third party to a fourth party.

The African Union adopted the progressive Convention on Cyber Security and Personal Data protection in 2014, although only 2 countries have ratified the convention. The Asia-Pacific Economic Cooperation Privacy Framework aims to develop uniform standard of data protection law across the region, while the APEC Cross-Border Privacy Rules (CBPR) system has been forged out of this framework but, unlike GDPR, it has no effect on domestic laws and regulations.[131]

6.4.3 Artificial Intelligence and Machine Learning

Data extraction and the surfeit of data,[132] alongside improvements in analytical capabilities using AI and ML[133] tools means that more analytics and decision-making is concentrated in algorithmic determinations.[134] Pools of data – that is, 'data lakes' - may be used for risk assessment in both regtech contexts as well as, for example, for decision-making on the viability of a customer or groups of customers.[135] The latter raises the potential for negative social outcomes. For example, as has been demonstrated in Kenya, fintechs have made lending decisions using exhaust data to extend increasing amounts of credit to persons who, *ab initio*, demonstrably cannot afford repayments as which has led to negative social outcomes.

Similarly, AI may result in inadvertently 'red-lining' of customer groups based on unknown algorithmic biases or even incorrect coding of AI algorithms.

This is the most public of the AI artefacts, which has led to increased regulatory scrutiny on AI processes. The need for increased 'explainability' – a term of art the AI/ML industry use for internal and regulatory reporting – from entities to regulators has correspondingly increased. Invariably, the industry undertakes 'model management' of AI as part of a risk management processes.[136] This means they combine data from various sources, undertake benchmarking and ensure that models have not drifted off from original designs and that internal governance is in place. Calibration takes place against recent historical data, and not using data that may have been subject to different (and now redundant) regulatory regimes. Similarly, the need for disclose to consumers that they may be interacting with an AI-powered 'bot' – such as a chatbot - is an emerging regulatory focus.[137]

For regulatory explainability, a privacy layer is usually inserted in the AI software stack by an entity being supervised to ensure conformity with privacy and anti-algorithmic bias rules and to allow their regulator to see data rather than the regulator having to subpoena all data. Many entities may use a model and data that has already been used to explain to regulators or build AI models that express interpretability on top of inputs of the AI model. These include scoring algorithms that inject 'noise' with what is known as local interpretable model-agnostic explanations (LIME) into their systems.

Building data pipes and maintaining privacy for use in AI within a bank is difficult: that's why it is outsourced. But this use of different vendors then creates a coordination issue in as some components are held by different vendors. The vendors then need to coordinate, including those storing the data and those modelling the data. Newer AI systems recognize that there may be a big restriction on data, and do not necessarily need to need to share all data to get a (validated) result based on AI processing and analysis. To validate the data though, some use Zero Knowledge Proofs[138] which allows sharing of data without revealing the underlying data subjects. As another method, often banks will use other entity's data in an aggregated form for data modelling and ML, which is often a method that avoids violate regulatory restrictions on data privacy, cloud computing and data localization.

6.4.4 Data localization and Cloud Computing

Many regulators have some rules around data localization, the scheme as to whether data sets relating to residents of a particular country should be stored on servers of an entity in its country of the entity's primary jurisdiction, and what data can be stored.

The entry of 'cloud computing' changes this paradigm, both at a localization level and what data is stored and used. Cloud computing-related regulations may be the domain of many regulators – for example a central bank, a telecommunications regulator and privacy regulator. A measure of regulatory coordination to enable cogent ancillary regulations is needed in furtherance of fintech and regtech policies.

6.4.5 Open Banking

6.4.5.1 Overview

Open banking frameworks vary in scope and requirements, ranging from prescriptive rules, to supervised facilitation, or just left to the market.

Prescriptive: This requires banks to share customer-permissioned data and requires third parties that want to access such data to register with particular regulatory or supervisory authorities. The scope and degree of prescription varies. The EU's PSD2, for example, applies only to specific types of data, like payments processing data, and provides third parties with both 'read' and 'write' access to data and payment initiation. The PSD2, though, does not prevent member jurisdictions from adopting a broader scope. Thereto, the UK also requires the nine largest banks and building societies to share publicly available information about branch and ATM locations, services and fees. In the UK, the Competition and Markets Authority (CMA) established Open Banking Implementation Entity (OBIE) to create software standards and industry guidelines for open banking. Australia's framework provides "read-only" rights for data aggregation purposes and will eventually cover industries beyond banking, such as the

telecommunications and energy sectors such that data can be shared across sectors.[139] The regulators involved in open banking vary too. In Australia, competition authorities are responsible for the implementation of open banking frameworks, while in the EU, India, Hong Kong and Singapore, the central bank or bank supervisor oversees the framework.[140]

Facilitated Supervision: Here regulators issue guidance and recommend standards, as well as release open API standards and technical specifications. This market scenario exists in Brazil, Mexico, and Singapore. While Singapore has provided guidelines, it has not imposed regulations on its leading banks. Driven by market adoption, numerous Singapore banks have opened their APIs. In 2018, Mexico passed its 'FinTech Law' ostensibly to cultivate an open banking standard that considers financial inclusion. In the US, the Consumer Financial Protection Bureau (CFPB) published principles on Principles For Consumer-Authorized Financial Data Sharing and Aggregation ostensibly to encourage competition, promote financial inclusion and protect consumers.[141]

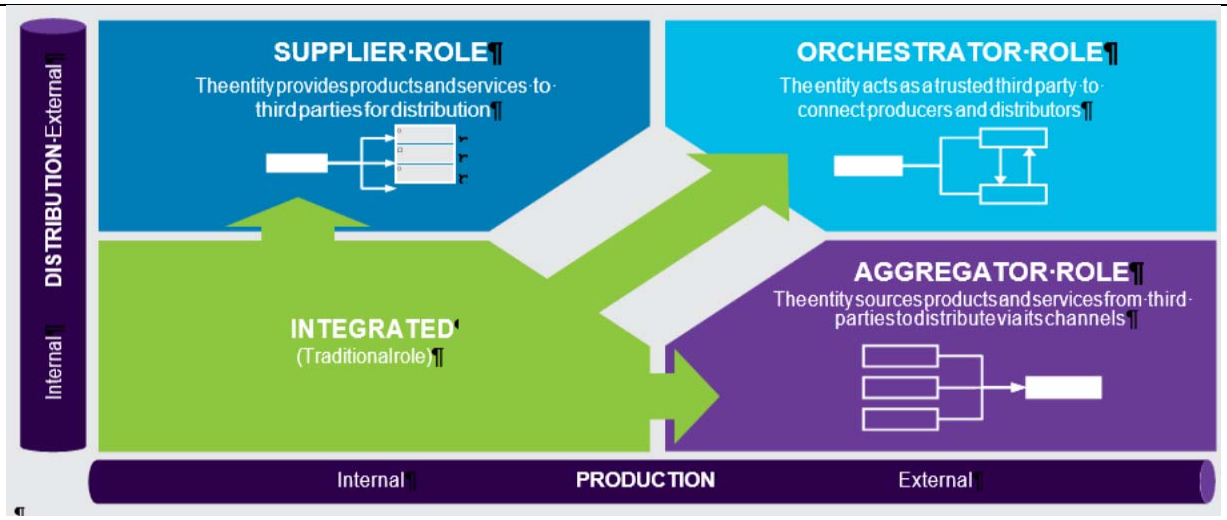
Market-driven approach: This approach has no explicit rules or guidance that require or prohibit the sharing of customer-permissioned data by banks with third parties.[142] There are benefits and challenges with each approach to open banking when balancing bank safety and soundness, encouraging innovation and consumer protection.[143] Here banks and fintechs can openly share data for slow, but steady, growth where large banks in these markets become aggregators and collaborate with fintechs to bring innovative offerings to their own existing customers – for example, in relation in budgeting-type applications for customers that undertake analysis of their spending habits. Often these collaborative initiatives are led by small regional banks who work with fintechs to take their products to larger and underbanked customer segments.[144] In India, the Unified Payments Interface (UPI) – a peer-to-peer payments scheme that bypasses intermediaries – has been widely adopted, while in China, fintech giants such as Ant Financial and Tencent are leveraging open APIs to allow third parties to offer services to their customers and make data more portable within their ecosystems.[145] **Exhibit 8** outlines the potential realignment of roles in a data-driven financial ecosystem using these open banking precepts.

6.4.5.2 API Development and Use

Some of the challenges hindering the development of APIs to share customer-permissioned data include the time and cost to build and maintain APIs and the lack of commonly accepted API standards. Some jurisdictions, such as Hong Kong and Singapore, issued recommendations on open API designs and technical specifications, aiming to facilitate adoption of open banking practices. As noted by the BIS, in jurisdictions where screen scraping or reverse engineering is still prevalent, banks are challenged with balancing security against ease of access. Banks generally prefer, or in some jurisdictions, are required to use more secure methods for sharing data for certain types of accounts, such as tokenized authentication through APIs, as opposed to screen scraping or reverse engineering.

Open Banking appears to be the start of a more structured collaborative environment in financial services, powered by enabling regulations and technology innovations such as API standardization and shared customer data insights. This is thought to be preceded by the revaluation and reassignment of traditional roles in financial services using an integrated marketplace, with specialized roles for each player.

In one vision developed by Cap Gemini and dubbed Open X, companies will assume a role in the financial sector that aligns with their capabilities and external operating environment. An 'integrated' role is the traditional role in which a company like a bank maintains full control of product and service creation as well as product distribution. The new paradigm envisages that ecosystem participants may instead assume the role of a supplier, aggregator or orchestrator. These roles are not business-model exclusive, but business-case specific and each ecosystem entity may mix and match roles depending on the business model.



In the 'supplier' role, the entity focuses on developing products and services, leaving distribution to a third-party or external player. In the 'aggregator' role, the entity delegates product and service creation to a third party or external entities but uses its internal channels for distribution. In the 'orchestrator' role, an entity plays a central role in coalescing ecosystem partners through connecting and coordinating their interactions to create the most value. This role links suppliers and aggregators and orchestrates their interactions. The new paradigm, it is suggested, means that an 'integrated' entity may struggle to match competitors' time to market and agility to quickly meet customers' personalized demands.

Exhibit 8: Potential realignment of roles in a data-driven financial ecosystem. Image Source: Cap Gemini Financial Services Analysis, 2019.

These secure methods enable banks to exercise greater control over the type and extent of data shared and enable more secure access management and monitoring. Several jurisdictions are issuing guidance on user authentication based on open API frameworks that require the use of tokenized protocols such as OAuth 2.020[146] open APIs, which will assist industry in transitioning away from screen scraping.[147] Some industry participants are looking at two potential monetization models for APIs - revenue-sharing and API access fees.[148]

6.4.5.3 Data access and transmission

Data access and transmission by third parties can range from a basic copy and paste screen scraping process to the transmission of standardized data elements using APIs. Despite broad emphasis on the importance of ensuring the security of customer-permissioned data, approaches for data access and transmission vary across jurisdictions according to respective legal and regulatory frameworks.[149]

In jurisdictions without explicit regulatory requirements, banks and third parties have more flexibility in data access and transmission practices. In these jurisdictions, the scope and process of data sharing may be governed by a contract executed between the bank and the third party.[150]

A number of jurisdictions have, or are in the process of developing, rules requiring disclosure and/or customer consent, but do not necessarily prescribe the exact contents of the disclosure form. Disclosure and consent requirements are primarily observed in contractual agreements between the banks and third parties.[151]

6.4.5.4 Screen Scraping

Most jurisdictions have no specific laws or regulations regarding the practice of screen scraping. In the US, courts have allowed screen scraping to continue where customer data or profiles are publicly available.[152] The EU's PSD2, regulates delivery of services requiring access to the user's account, introducing new types of payment services: payment initiation services, account information services, and the service of confirmation of the availability of funds in a payment account. The PSD2 requires 'account servicing payment service providers' (such as banks) to ensure access and prepare an interface for providers of these new payment services.

The EU PSD2, though, limits third parties from screen scraping for payment account data through banks' standard customer interface. Banks instead either offer dedicated APIs or a modified customer interface that enables third parties to identify themselves using authentication certificates when accessing customer data. Third parties use screen scraping techniques from this modified user interface, but the interface may limit or control the data available to the third party. This modified customer interface would also be used as a contingency mechanism when the bank's API is unavailable. In the EU, banks can be exempted from setting up a contingency mechanism if their competent authorities determine that the bank's dedicated APIs comply with certain conditions.

The technical conditions for access by these entities are defined in Commission Delegated Regulation (EU) 2018/389 on regulatory technical standards for strong customer authentication and common and secure open standards of communication known as regulatory technical standards (RTS). Providing access to existing user interfaces offers a 'backdoor' enabling the use of screen scraping in a new, modified form. Because the earlier method of screen scraping must be appropriately modified and equipped with new functions (the obligation for authentication of the Account Servicing Payment Service Provider (ASPSP), the new version allowed by the RTS Regulation is often referred to as 'screen scraping plus.'[153] The additional requirements established for this method are intended to eliminate the defects that previously disqualified this method in the view of regulators.[154]

However, since PSD2 is only about payment accounts, nothing in PSD2 regulates non-payment accounts - including access.[155] Similarly, there are no obligations in any part of the legislation for ASPSPs to identify third party providers (TPPs) such that http headers' signature mechanism does not require any collaboration from the ASPSP, neither from a technical point of view nor from a legal perspective.[156]

6.4.6 DLT Regulation

A sample of the legal issues that would appear to be most pertinent to DLTs include the legality and enforceability of smart contracts; evidential weight of DLT-derived data; property rights in crypto-assets; time and place of contracting using a blockchain and smart contracts; the 'chain' of legal liabilities in the sector; competition issues in a decentralized environment; criminal use and liability; and which court may have jurisdiction over a matter involving DLTs and their applications in a 'distributed' multi-national nodes environment. As DLT-related regulations may be the domain of many regulators – for example a central bank, and telecommunications regulator, and privacy regulator – a measure of regulatory coordination to enable cogent ancillary regulations is needed in furtherance of fintech and regtech policies.

6.4.7 Risk Management and Liability

With more parties and intermediaries involved in the provision of financial services in an open banking model, risks increase, especially for banks who provide the data to third parties. In some jurisdictions, outsourcing policies place responsibility on banks to ensure third parties are compliant with these rules, and generally stipulate documentation as part of contractual arrangements. In other jurisdictions, bank supervisors have supervisory authority over registered third parties. Data sharing, storage and security requirements, apply mostly to banks and outsourced bank services, and not necessarily with the third parties contracting directly with bank customers. Indeed, bank regulators have limited authority especially over those that are not registered with a separate authority.[157] Even in jurisdictions where there are established liability rules, Banks may face reputational risk.

6.4.8 Consumer Protection

A number of risks for users of open banking ecosystems include data breaches that can lead to identity theft, and subsequent financial losses for customers; unauthorized payments or transactions made without the account holder's permission if log-in credentials are accessed by untrusted parties or from errors in (or attacks to) the functioning of payment initiation services; and defective payments or transactions, requested by the customer but wrongly processed by the providers involved.[158]

But with more parties and intermediaries involved in the provision of financial services in an open banking model, it is more difficult to assign liability and the amount of damages to the customer, if any, in the event of financial loss, or where there has been erroneous sharing or loss of sensitive data. Additionally, consumer protection laws may not have been updated to take open banking business models into consideration.

6.4.9 Cybersecurity

While there are manifest benefits – particularly to competition – in an open banking regime, with customer data travelling a complex supply chain, it also has the potential to magnify the impact of breach and cybersecurity incidents. The newly enabled third parties accessing data in an open banking regime sit outside the perimeter of bank security, and banks will be interacting with them without clear understanding of their system's security posture. Thus, banks may be exposed to new threats emanating from beyond their traditional areas of control. Data collected by third parties, whether via screen scraping, reverse engineering or tokenized authentication methods through APIs, can be stolen or compromised. Data can become compromised during transit, at-rest (storage) or in-use.

The EU outlines its data storage and security requirements for data sharing under PSD2's RTS. Where the third party is authorized. EU banks generally are not expected to inspect or monitor the data security frameworks put in place by the authorized third party. The UK has adopted a common authentication protocol called OAuth 2.0. to provide a secure method for verifying digital identities and provides a formal structure using tokens for obtaining, and securely transferring, consumer consent between entities.[159] The Digital ID & Authentication Council of Canada recently released its 'Pan-Canadian Trust Framework,' which may form part of a secure open banking framework.[160]

6.5 Potential Regulatory Gaps in Data Protection

While there is a trend to balance access to data with data protection rules, a number of gaps remain or are being caused by new rules. In many cases, the laws do not cover all use cases, nor make provisions for technology advances that – in effect – facilitate regulatory dialectics.

Many national data protection laws contain significant gaps and exemptions, identified by the BIS as:[161]

- Exclusion for small businesses (Australia and Canada)
- Small data sets (for example, Japan which excludes data sets with less than 5,000 entries)
- Types of data subject (e.g. only to children, or not to employee data)
- Sensitivity of data (e.g. only to sensitive data like health or financial records)
- Sources of data (e.g. restricted to either online or offline data collection)
- Sectorial data (e.g. exemptions related to the private and public sector, or laws that are restricted to specific sectors like health and credit)
- Allowing individual companies to determine the 'scope' of the data protection that they offer to consumers[162]

Data protection laws worldwide do not consider anonymous data as being personal data, allowing it to be freely used and shared. However, advances in technology means that there is also the potential for a privacy-shattering de-anonymization of pools of these de-anonymized data sets. As one study[163] showed, that once bought, the data can often be reverse engineered using machine learning to re-identify individuals, despite the anonymization techniques. The gap occurs in so far as the de-anonymized data is no longer subject to data protection regulations, so it can be freely used and sold to third parties like advertising companies and data brokers.[164] Rules may be

required to design training to be much more than simply adding noise, sampling datasets, and other de-identification techniques.

The GDPR and the California Consumer Privacy Act consider that each and every person in a dataset has to be protected for the dataset to be considered anonymous. The GDPR in particular brings into scope 'pseudonymous data' – data that does not contain obvious identifiers but might yet be re-identifiable[165] by explicitly adding references to pseudonymisation as an intermediate form of de-identification.

7 Conclusions

Given though the variety of authorities involved, some degree of regulatory coordination may be needed to address potential regulatory inconsistencies or gaps.

The result is that which identifies specific persons. This is often an artifact of the process of using data derived from multiple data sets when 'training' AI systems through ML for making automated decisions and forecasts. Regulatory gaps occur when data is 'sampled' and anonymized, a process which includes removing identifying characteristics in data sets such as names and email addresses such that, in theory, there is no ability to uniquely identity individuals.

Fintech moves traditional processes to new entrants – fintechs – and new methodologies, with new technologies and business processes at the core of many of the offerings. Similarly, regtech provides a more streamlined ability to undertake compliance. The emergence of fintech, fintechs, and regtech, we show, intersect at the storage and use of data, be that through cloud computing, DLTs, AI/ML, and open banking APIs.

Enabling fintechs, and then regulating fintech and the surfeit of data at the core of fintech and regtech, creates regulatory challenges. With common technology components driving fintech and regtech, this provides a catalyst for a good time to reengineer regulatory methodology.

There has, however, been little analysis, so far, as to how a comprehensive regtech and fintech ecosystem for data-driven finance could and should be developed in a given financial system that also embraces a regulatory framework that nurtures and catalyzes these innovations. Given the integrated nature of fintech and regtech, any new framework or policy, we show, would necessarily need to include ancillary regulations that affect both fintech and regtech, such as relating to AML, AI/ML, cloud computing, data use and protection, distributed ledger technology, and cyber-security.

Similarly, where the power is in the data, a new systemic risk may arise from concentration of data in the hands of relatively fewer technology firms which may replace and complement (financial) systemic risk represented by banks that were too-big-too-fail or too connected-too-fail. This may be ameliorated though through use of open APIs where there is either mandated access data by fintechs for both fintech and regtech use, the latter in the form of centralized eKYC solutions.

This study demonstrates though that, beyond questions of the interaction between financial and data regulation, are questions around the role of technology in regulation, compliance and digital financial transformation. That is, the role of fintech, fintechs and regtech both in supporting the process of transition and providing the basis of a system to address its requirements, monitor compliance and support the achievement of regulatory and policy objectives by regulators and policymakers.

It is important to note that to support this approach and to avoid regulatory ambiguity, gaps and arbitrage, regulations (as needed) for ancillary-type services – such as cloud computing and AI – would need to be developed, or existing regulations clarified to allow authorized/licensed fintechs to provide financial services where these services, functionally, touch on handling or storing customer funds and on AML.

Noting the above, we apply the Occam Razor Principle - the eponymous problem-solving principle that the simplest solution tends to be the best – to the contours of an appropriate regulatory strategy for fintech and regtech.

Thereto, we propose an evolutionary, principles-based functional approach to fintech regulation and regtech development that also supports related ancillary services around data accumulation and use. Regulatory actions may vary from "disclosure" to "light-touch regulation and supervision" to a "full-fledged regulation and supervision", depending on the risk implications. A tiered approach could be use, increasing oversight as an entity grows and its risk profile changes.

This may include implementation of varieties of regulatory sandboxes to act as a buffer – in a codified 'transition period' – if regulators wish to move from the strict rules-based, institutional approach to the more flexible and encompassing regime. This transition also avoids a 'big bang' approach which some regulators would find too overwhelming, given their need to retain elements – particularly for the banking and lending sector – of the rules-based, institutional approach.

These approaches, it is submitted, would satisfy the following:

- **Regulators,** as there is no 'bang-bang' move from a familiar and tested institutional/product/rules-based approach towards a potentially unfamiliar principles/functional/product approach that may raise public-policy and concerns and potentially even legal challenges. It also allows regulators must also act in the interests of customers, protecting them in a changing environment that can pose new, unanticipated risks that may also raise systemic stability and AML concerns. It also allows multiple regulators to have oversight on sectors implementing specific functions/products, without significantly impacting their respective remits and creating regulatory arbitrage. Regtech solutions would also improve supervisory capabilities and compliance by fintechs.

- **Fintechs (as SMEs),** as it provides an opening to introduce into the market innovations with less regulatory requirements. Regtech solutions would also improve internal risk management capabilities and regulatory compliance.

- **Plans for Sandboxes,** as this glacial approach fits within a risk-based approach that cultivates innovative classes or services and products whilst limiting the potential of open-ended regulation to exposure, ML and consumer harm.

Similarly and critically, ancillary regulation that would be a touch-point of both regtech and fintech ecosystem would be needed to close any potential regulatory gaps, ensuring regulatory certainly in the use of technologies and the surfeit of data powering both fintechs and regtech.

Noting the above, we apply the Occam Razor Principle - the eponymous problem-solving principle that the simplest solution tends to be the best – to the contours of an appropriate regulatory strategy for fintech and regtech. Thereto, we propose an evolutionary, principles-based functional approach to fintech regulation and regtech development that also supports related ancillary services around data accumulation and use.

In essence, these ancillary regulations, as needed, would address the intersection of fintech and regtech in use of data sets. This could relate to use of personal data; cloud computing and data localization/safe harbor rules; sharing of data for anti-money laundering purposes; rules around recognizing data stored on DLT/blockchain for evidential and other purposes. This also includes recognizing the growing use of AI and ML to analyze with calibrated models in a manner that does not create or perpetuate algorithmic biases and unintended red-lining of classes of people for access to financial services and products.

[1] Leon Perlman Ph.D.; Head: Digital Financial Services Observatory, Columbia Institute for Tele-information, Columbia University (CITI), New York.

[2] This paper was supported by a grant by the Bill and Melinda Gates Foundation. I am grateful to Professor Eli Noam, Director of CITI, for his supportive role and comments on early drafts; and to my colleagues at the Digital Financial Services Observatory.

[3] Zetzsche, D; Buckley, R; Arner, D and Weber, R (2019) T*he Future of Data-Driven Finance and Regtech: Lessons from EU Big Bang II*, available at https://ssrn.com/abstract=3359399; Weber, R (2017) Regtech as A New Legal Challenge, available at: https://ssrn.com/abstract=3359399

[4] Armstrong, P (2017 *Regulatory Technology: Reshaping the Supervisor-Market Participant Relationship*, available at https://bit.ly/34HZXPs

[5] di Castri, Simone and Grasser, Matt and Kulenkampff, Arend (2018) *Financial Authorities in the Era of Data Abundance: Regtech for Regulators and Suptech Solution*s, available at https://ssrn.com/abstract=3249283

[6] FSB (2019) *FinTech and market structure in financial services: Market developments and potential financial stability implications*, available at https://www.fsb.org/2019/02/fintech-and-market-structure-in-financial-services-market-developments-and-potential-financial-stability-implications/

[7] Basel Committee on Banking Supervision (2013) *Principles for Effective Risk Data Aggregation and Risk Reporting,* available at https://www.bis.org/publ/bcbs239.pdf

[8] Sweezy, P (1943) *Professor Schumpeter's Theory of Innovation*, available at https://www.jstor.org/stable/19245; and Stewart, L (2010) *The Impact of Regulation on Innovation in the United States: A Cross-Industry Literature Review*, available at https://www.itif.org/files/2011-impact-regulation-innovation.pdf. Clearly a firm can innovate without ever inventing. See also

[9] Zetzsche, D; Buckley, R; Arner, D and Weber, R (2019) *The Future of Data-Driven Finance and Regtech: Lessons from EU Big Bang II*, available at https://ssrn.com/abstract=3359399; Weber, R (2017) Regtech as A New Legal Challenge, 46 J. FIN. TRANSFORMATION 10 (2017). Electronic copy available at: https://ssrn.com/abstract=3359399

[10] The ongoing trade war between the US and China led global fintech investments to fall by 29% during 1H19 according to Accenture. During 2014 around USD 12 billion was invested in Fintech companies, and in 2015 USD 20 billion. TechRadar (2019) *Global fintech investment plummets worldwide*, available at https://www.techradar.com/news/global-fintech-investment-plummets-worldwide

[11] Zetzsche, D; Buckley, R; Arner, D and Weber, R (2019) *The Future of Data-Driven Finance and Regtech: Lessons from EU Big Bang II*, available at https://ssrn.com/abstract=3359399

[12] Humby is widely credited as the first to coin the phrase. Palmer, M (2006) *Data is the New Oil*, available at http://ana.blogs.com/maestros/2006/11/data_is_the_new.html

[13] Determann, L (2018) *No One Owns Data*, available at https://ssrn.com/abstract=3123957

[14] Baškarada, S & Koronios, A (2013) *Data, Information, Knowledge, Wisdom (DIKW): A Semiotic Theoretical and Empirical Exploration of the Hierarchy and Its Quality Dimension*, available at https://journal.acs.org.au/index.php/ajis/article/view/748

[15] Determann, L (2018) *No One Owns Data*, available at https://ssrn.com/abstract=3123957

[16] *ibid*

[17] Rubinstein, Ira (2012) *Big Data: The End of Privacy or a New Beginning?*, available at https://bit.ly/34yLcyt

[18] *ibid*

[19] Stucke, M & Grunes, A (2016) *Introduction: Big Data and Competition Policy*, available at https://ssrn.com/abstract=2849074

[20] Stucke, M & Grunes, A (2016) *Introduction: Big Data and Competition Policy*, available at https://ssrn.com/abstract=2849074. See also BIS (2018) *Sound Practices: Implications of fintech developments for banks and bank supervisors*, available at https://bit.ly/34G0Jwr

[21] For an introduction to DFS, see Perlman, L (2019) *Introduction to Digital financial Services,* available at www.dfsobservatoiry.com

[22] See further on the nature of adverse selection and data sets, Mazer & Rowan (2016) *ibid;* and generally on big data and DFS, Chen, G & Faz, X (2015) *The Potential of Digital Data: How Far Can It Advance Financial Inclusion*?, available at https://goo.gl/dxxSIU

[23] This information asymmetry, in a credit-provision context, may result in what is termed adverse selection, such that without a credit risk assessment – or credit score – the borrower will seek and often be given credit by lenders who are unable to obtain enough information on hand to have made a more seasoned determination of whether the loan would be repaid. Thus, those with access to cogent data sets will mitigate the risks of adverse selection. See further Mazer & Rowan (2016) *ibid*

[24] Christensen P (2015) *Credit Where Credit Is Due*, available at https://goo.gl/h0Oapm

[25] Data can of course be gleaned from bank-related activity but this may be restricted through bank secrecy laws in some countries - for example Pakistan - which have often gotten in the way of sharing data that could otherwise be valuable in the hands of alternative financial providers. Here then, traditional credit providers benefit from their 'proprietary' data.

[26] See San Pedro, J *et al* (2015) *MobiScore: Towards Universal Credit Scoring from Mobile Phone Data,* available from https://goo.gl/Mkwp5T

[27] Also through some feature phones that have Facebook, Twitter, and Whatsapp installed. See further, Perlman, L (2017) *DFS Handset Overview: ITU FG on DFS*, available at http://www.itu.int/en/ITU-T/focusgroups/dfs/Pages/default.aspx

[28] In most cases prospective (and existing) users can only install and thus the app to get credit only if they agree to all these metrics being monitored.

[29] See Government Of Kenya (2016) *Gazette Notice No. 678: Proposed Market Inquiry And Sector Study On The Kenya Banking Sector-Phase II By Competition Authority Of Kenya,* available at https://goo.gl/wbqDX6

[30] Newtech.law (2019) *Can a user's account be accessed through screen scraping?*, available at https://bit.ly/2sy2w8J

[31] American Banker (2017) *Fintechs' defense of screen scraping is shortsighted*, available at https://bit.ly/33AOKPl

[32] There are a number of API types. An Open API is an interface that provides a means of accessing data based on a public standard. Also known as external or public API; Internal/Closed API is an interface that provides a means of accessing data based on a private standard. Also known as internal API; Partner API is an API created with one or two strategic partners who will create applications, add-ons, or integrations with the API. BIS (2019) *Report on open banking and application programming* interfaces, available at https://www.bis.org/bcbs/publ/d486.htm

[33] Data related to payment services is more commonly accessed via APIs, while data for information purposes (such as balances and transaction histories) are commonly accessed via screen scraping. This phenomenon may also be due to regulatory requirements in some jurisdictions, such as in the EU where payments-related data is expected to be shared via APIs. Pallardó, A (2016) *PSD2: Screen Scraping vs APIs?*, available at https://www.kantox.com/en/psd2-screen-scraping-vs-apis/, and BIS (2019)

[34] What makes AI systems so powerful is that it is a different logic to that of humans who will always ask 'why' and seek correlation with data sets. With AI, correlation between data sets is the conclusion it derives at. The AI dilemma for entities is whether to use the (output) data knowing humans may not necessarily understand the methodology – including ML - used to create it.

[35] Linear regression is what your models use for typical model building, everything else may be supervised learning where no additional input is needed. Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs. It infers a function from labeled training data consisting of a set of training examples.

[36] Fraud detection is moving to real-time monitoring. That is, if an entity is doing real-time monitoring, then they do not need to hold on to data. Credit card determinations and applications are based on multiple models, from regression to now machine-learning models. Algorithms used in real-time are not just memorizing anomaly detection but now generalizing them.

[37] Kaminski, M (2019) *The Right to Explanation, Explained*, available https://ssrn.com/abstract=3196985

[38] This is part of the Customer Data Right regulations granting greater access to consumer data. . In Singapore, open banking is regulated by a non-mandatory governance framework. See also Bahrain FinTech Bay (2019) *FinTech Regulations Report 2019,* available at https://www.fintech-consortium.com/single-post/2019/02/25/Bahrain-FinTech-Bay-Launches-Bahrain-FinTech-Regulations-Report-2019

[39] Chappelow, J (2019) Open Banking, available at https://bit.ly/2pqXJVh

[40] European Banking Authority (2018) *Risks And Opportunities Arising From Fintech*, available at: https://ssrn.com/abstract=3359399

[41] Bankrate UK (2019) Which banks support open banking today?, available at https://www.bankrate.com/uk/open-banking/which-banks-support-open-banking-today/
[42] BIS (2019)
[43] BIS (2019) and below.
[44] *ibid*
[45] WIRED UK (2018) *What is Open Banking and PSD2? WIRED explains*, available at https://www.wired.co.uk/article/open-banking-cma-psd2-explained
[46] If for example a third party (or bank) misuses customer data, then the case may be subject to data protection laws and action from regulators.
[47] Stucke, M & Grunes, A (2016) *Introduction: Big Data and Competition Policy*, available at https://ssrn.com/abstract=2849074
[48] Access Now (2018) *Data protection: why it matters and how to protect it*, available at https://www.accessnow.org/data-protection-matters-protect/
[49] Schwartz, & Solove, D (2013) Reconciling Personal Information in the United States and European Union, available at SSRN: https://ssrn.com/abstract=2271442
[50] OECD (2013) Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, available at https://bit.ly/2RaDy9A
[51] Council Regulation 2016/679, *supra* note 13, arts. 13–15
[52] Kaminski, M (2019), *Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability*, available at https://ssrn.com/abstract=3351404
[53] EU (2019) *What does data protection 'by design' and 'by default' mean?,* available at https://bit.ly/34D9XcZ
[54] *ibid*
[55] On use of DLTs in finance, see Perlman, L (2019) *Regulation of the Crypto-economy*, available at www.dfsobservatory.com
[56] The emergence thought of permissioned, controlled 'consortia' DLTs for use in banking and other verticals has altered this 'totally decentralized' paradigm.
[57] Mitra, R (2019) *What is Web 3.0? The Evolution of the Internet*, available at https://blockgeeks.com/guides/web-3-0/
[58] See Perlman, L (2019) *Regulation of the Crypto-economy,* available at www.dfsobservatory.com
[59] McKinsey (2018) *How secure is the global financial system a decade after the crisis?,* available at https://mck.co/2OBQO35
[60] Arner, D, Barberis, J & Buckley, R (2017) *FinTech and Regtech In A Nutshell, And The Future In A Sandbox*, available at https://cfa.is/2POCyVl
[61] Cesa, M (2017) *A Brief History of Quantitative Finance,* available at https://bit.ly/2JFkN8b; Dincer, H & Hacioglu, U (2014) *Globalization of Financial Institutions: A Competitive Approach to Finance,* available at https://bit.ly/2rjxQnv; Celik, H (2013) *The Impacts of Information Technologies on Financial Institutions,* available at https://bit.ly/2GSm8Xq; Arner, D, Barberis, J & Buckley, R (2017) *ibid.*
[62] Between 2008 and 2015, there has been a 492% increase in annual volume of regulatory changes. Kocianski, S (2016). *The Regtech Report: Global Regulatory Requirements Are Creating a Huge Opportunity for Regtech Firms*, available at https://goo.gl/YF1paV
[63]Development ASIA (2017) *How Regtech is Helping Banks Manage Risks*, available at https://bit.ly/2CzOC9c
[64] Schutzer, D (2017) *Regtech: Innovation and The Future of Financial Services*, available at https://bit.ly/2udvar1
[65] Based on conversation with RBI, NRB, CNVB, Mann, P (2017) *Regtech: The Emergence of the Next Big Disruptor*, available at https://bit.ly/2gFm2XL
[66] Microsoft excel was introduced for Macintosh in 1985 and for Windows in 1987.
[67] XML, Extensive Markup Language, is a computer language that allow users to create self-describing data using tags, elements and attributes. It helps simplify data interchange, enable smart code as well as smart searches. Tidwell, D (2002) *Introduction to XML,* available at https://ibm.co/2EFSPFq
[68] FCA (2016) *Feedback Statement Call for Input on Supporting the Development and Adopters of Regtech*, available at https://bit.ly/2bXLSrg
[69] Companies leveraging their knowledge of technology and data from their primary business to improve existing financial system processes and capabilities. This may be tech or e-commerce companies that are already connected to large number of clients and hence contains large volumes of data. Summarized from Arner, D, Barberis, J,

Buckley, R, et al. (2017) *From FinTech to TechFin: The Regulatory Challenges of Data-Driven Finance,* available at https://bit.ly/2HgS0bq; Shea, R (2016) *Fintech Versus Techfin: Does Technology Offer Real Innovation Or Simply Improve What Is Out There?,* available at https://tmsnrt.rs/2GROwwJ

[70] Infosys (2017) *Fintech Revolution in Banking: Leading The Way to Digital, a*vailable at https://infy.com/2GXyjly; Robles, P (2016) *Five Ways Fintech Upstarts Are Disrupting Established Financial Institutions*, available at https://bit.ly/2aw5ukg; Crook, R (2017) *The Race Is On to Disrupt Traditional Banking,* available at https://bit.ly/2GRbUdX

[71]Accenture (2018) *The Growing Need for Regtech,* available at https://accntu.re/2FG4jcC, Schutzer, D (2017) *Regtech: Innovation and the Future of Financial Serv*ices, available at https://bit.ly/2udvar1

[72] Examples include AI, API, Big Data Analytics, Biometrics, Cloud computing, and DLT.

[73] Conforming to a rule- law, regulation, and standard. Regulatory compliance is a set of guidelines that a business must follow as per the requirements set by regulators. Summarized from International Compliance Association (2018) *What is Compliance?,* available at https://bit.ly/2iNpZbj; The University of Scranton (2018) *A Definition on Regulatory Compliance,* available at https://bit.ly/2wa6w0v

[74] Any person, other than a representative, who regularly furnishes advices, renders intermediary services or both. For example, Financial Institutions and FinTechs. FSB (2018) *FAIS – Understanding the Practicalities*, available at https://bit.ly/2yvf3Kb

[75] See Arner, D, Barberis, J & Buckley, R (2017) *FinTech and Regtech in a Nutshell, and the Future in a Sandbox*, available at https://cfa.is/2POCyVl

[76] IFC (2016) *De-risking by Banks in Emerging Markets- Effects and Responses for Trade*, available at https://bit.ly/2JGkj1p

[77] The Board includes all G20 major economies, Financial Stability Forum members, and the European Commission. It is hosted and funded by the BIS and is based in Switzerland. See FSB (2018) *Our History*, available at http://www.fsb.org/about/history/

[78] Eyers, J (2016) *Welcome to the new world of 'regtech'*, available at https://bit.ly/2dAH5Dz

[79] On average, 201 regulatory alerts daily. Hill, E (2016) *Is Regtech The Answer To The Rising Cost Of Compliance?*, available at https://bit.ly/2q6CBAn

[80] CB Insights (2018) *The Evolution of Regtech and the Future of Regulatory Compliance*, available at https://bit.ly/2EH9hG7

[81] See IFC (2016) *De-risking by Banks in Emerging Markets- Effects and Responses for Trade,* available at https://bit.ly/2JGkj1p

[82] Chappell, L & Dolphin, T (2010) *The Effect of the Global Financial Crisis on Emerging and Developing Economies*, available at https://bit.ly/2qpO6CQ

[83] IFC (2017) *Digital Financial Services: Challenges and Opportunities for Emerging Market Banks*, available at https://bit.ly/2wAEqq8

[84] CGAP (2018) *Digital Financial Services*, available at https://bit.ly/1RrXkXY

[85] See IFC (2017) *Digital Financial Services: Challenges and Opportunities for Emerging Market Banks*, available at https://bit.ly/2wAEqq8; USAID (2016) *Understanding the Risks of Digital Financial Services,* available at https://bit.ly/2IQW1Ro; Arenaza, S (2014) *Potential Risks To Clients When Using Digital Financial Services,* available at https://bit.ly/1z8W5AX

[86] Toronto Center (2017) *FinTech, Regtech and SupTech: What They Mean for Financial Supervision*, available at https://goo.gl/R3vWxH

[87] Around 10-15% of the workforce is dedicated to governance, risk management and compliance. Deloitte (2017) *The Regtech Universe On the Rise*, available at https://goo.gl/LUpKtH

[88] See Petrasic, K (2016) *Regtech Rising: Automating Regulation for Financial Institutions,* available at https://bit.ly/2rhI1cL

[89] Standards to manage money laundering and terrorist financing risks for financial institutions. Money laundering refers to the conversion or transfer of property or any association, knowing it is derived from criminal activity, for the purpose of hiding its origins, nature, location, disposition, movement, ownership. Similarly, financing of terrorism is the provision or collection of funds to contribute to the commission of specific offences while in complete knowledge that they are being used or will be used for such purposes. Summarized from European Investment Bank Group (2018) *Anti-Money Laundering and Combating Financing of Terrorism Framework,* available at https://bit.ly/2HX4QMH; CGAP (2005) *AML/CFT Regulation,* available at

http://www.cgap.org/publications/amlcft-regulation; IMF (2018) *Anti-Money Laundering/ Combating the Financing of Terrorism (AML/CFT)*, available at https://www.imf.org/external/np/leg/amlcft/eng/

[90] ACAMS defines KYC as: 'AML policies and procedures used to determine the true identity of a customer and the type of activity that is "normal and expected," and to detect activity that is "unusual" for a particular customer.' ACAMS (2018) *AML Glossary of Terms*, available at https://www.acams.org/aml-glossary/. Although FATF largely discarded the term 'KYC' in its documents onwards from 2003, KYC is still widely used and is now considered to be only but one – the identity input – component of a CIV procedure that, in turn, is part of the ongoing CDD process. Perlman, L & Gurung, N (2018) *The Use of eIDs and eKYC for Customer Identity and Verification in Developing Countries: Progress and Challenges,* available at www.dfsobservatory.com

[91] ACAMS defines CDD in terms of ML controls, as requiring 'policies, practices and procedures that enable a financial institution to predict with relative certainty the types of transactions in which the customer is likely to engage. CDD includes not only establishing the identity of customers, but also establishing a baseline of account activity to identify those transactions that do not conform to normal or expected transactions.' FATF (2012) *The FATF Recommendations,* available at https://bit.ly/1e7w0Gl. See Lyman, T & de Koker, L (2018) *KYC Utilities & Beyond: Solutions for AML/CFT Paradox?*, available at https://bit.ly/2OqOgso.

[92] PwC (2017) *Get Ready for Regtech,* available at https://pwc.to/2JSDZyH; Bryans, T (2017) *The Rise of Regtech and the Impact on Compliance*, available at http://www.corporatecomplianceinsights.com/the-rise-of-regtech/; Schutzer, D (2017) *Regtech: Innovation and the Future of Financial Servi*ces, available at https://bit.ly/2udvar1

[93] Wakefield, N (2017) *How Technology Is Driving Financial Inclusion*, available at https://bit.ly/2HuHxH3

[94] Older technologies related to previous or outdated systems. They can include reporting using excel and XML templates, emails, CDs, and paper by market participants to the central bank and recording and storing information in excel spreadsheets by central banks. Legacy technologies may allow for sufficient support to run central bank operations but may be insufficient to allow them to scale and adapt to the changing financial sector. Summarized based on conversation with RBI, NRB, CNVB, and information from Mann, P (2017) *Regtech: The Emergence of the Next Big Disruptor*, available at https://internationalbanker.com/finance/regtech-emergence-next-big-disruptor/; Schneider, A (2013) *When Companies Become Prisoners of Legacy Systems,* available at https://bit.ly/1LJYxZH

[95] Mann, P (2017) *Regtech: The Emergence of the Next Big Disruptor*, available at https://bit.ly/2gFm2XL; Deloitte (2017) *Regtech is the new FinTech,* available at https://bit.ly/2IetXui

[96] Dias, D & Staschen, S (2017) *Data Collection by Supervisors of Digital Financial Services,* available at https://bit.ly/2LCbTMB

[97] Arner, D, Barberis, J & Buckley, R (2017) *FinTech and Regtech in a Nutshell, and the Future in a Sandbox*, available at https://cfa.is/2POCyVl

[98] CNBV is an independent agency of the Secretariat of Finance and Public Credit body with technical autonomy and executive powers over the Mexican financial system

[99] Source: CNBV Website.

[100] Arner, D; Barberis, J and Buckley, R. (2017) *Fintech and Regtech in a Nutshell, and the Future in a Sandbox*, University of Hong Kong Faculty of Law Research Paper No. 2017/040, available at https://ssrn.com/abstract=3088303

[101] Arner, D; Barberis, J and Buckley, R ., (2016) *The Emergence of Regtech 2.0: From Know Your Customer to Know Your Data*, available at https://ssrn.com/abstract=3044280

[102] WEF (2016*) The Complex Regulatory Landscape for FinTech An Uncertain Future for Small and Medium-Sized Enterprise Lending*, available at https://bit.ly/2q7SQAO

[103] IMF (2017) *Fintech and Financial Services: Initial Considerations*, available at https://bit.ly/2suIUCe

[104] RBI (2018) *Report of the Working Group on FinTech and Digital Banking,* available at https://bit.ly/34G8xhT

[105] Stewart, L (2010) *The Impact of Regulation on Innovation in the United States: A Cross-Industry Literature Review,* available at https://www.itif.org/files/2011-impact-regulation-innovation.pdf

[106] European Banking Authority (2018) *Risks And Opportunities Arising From Fintech*, available at: https://ssrn.com/abstract=3359399

[107] RBI (2018) *Report of the Working Group on FinTech and Digital Banking*, available at https://bit.ly/2LeQp6Q

[108] Frantz, P & Instefjord, N, (2015) *Rules vs Principles Based Financial Regulation*, available at https://ssrn.com/abstract=2561370

[109] World bank (2018) *The Bali Fintech Agenda: A Blueprint for Successfully Harnessing Fintech's Opportunities,* available at https://www.worldbank.org/en/news/press-release/2018/10/11/bali-fintech-agenda-a-blueprint-for-successfully-harnessing-fintechs-opportunities
[110] BIS (2019)
[111] Zetzsche, D; Buckley, R; Arner, D & Barberis, J (2017) *Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation*, 23 FORDH. J. CORP. FIN. L. LAW 31-103
[112] BIS (2019)
[113] Enriques, L (2017) *Financial Supervisors and Regtech: Four Roles and Four Challenges*, available at https://ssrn.com/abstract=3087292
[114] Adapted from BIS (2019)
[115] Toronto Center (2017) *FinTech, Regtech and SupTech: What They Mean for Financial Supervision*, available at https://goo.gl/R3vWxH
[116] For example, the EU's GDPR 'right to be forgotten' requirements appear incompatible with data stored on blockchains due to the inability to delete data off a blockchain. For a discussion, see Renieris, E (2019) *Forget erasure: why blockchain is really incompatible with the* GDPR, available athttps://bit.ly/2P2ONyi. And Visa is launching a new blockchain-based system ostensibly to provide transparency and control to users by enforcing agreed-upon policies on what functions can be evaluated over private data even when the users are online, and enforcing the set of parties with whom the results are shared. AMBCrypto (2019) *Visa's R&D wing reveals new blockchain system dubbed 'LucidiTEE'*, available at https://bit.ly/35T9vr5
[117] *ibid.*
[118] CFTC 2018) *LabCFTC,* available at http://www.cftc.gov/LabCFTC/Overview/index.htm; Gach, R & Gotsch, M (2017) *Fintech Goes to Washington: Regulators, Financial Firms Discuss Wave Of Future,* available at https://bit.ly/2xRexEv; Loizos, C (2017) *Startups Say This Fintech 'Lab' Is Giving Them Needed Access To Wall Street And Regulators,* available at https://tcrn.ch/2qqV2jP
[119] Stucke, M & Grunes, A (2016) *Introduction: Big Data and Competition Policy*, available at https://ssrn.com/abstract=2849074
[120] Consumers International (2018) *The State Of Data Protection Rules Around The World A Briefing For Consumer Organizations,* available at https://www.consumersinternational.org/media/155133/gdpr-briefing.pdf; UNCTAD
[121] DLA Piper (2019)
[122] OECD (2013) *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, available at https://bit.ly/2L9MiJb
[123] Council Regulation 2016/679, *supra* note 13, arts. 13–15
[124] 15(1)(g), 15(3), at
[125] Kamara, I (2017) *Co-regulation in EU Personal Data Protection: The Case of Technical Standards and the Privacy by Design Standardisation 'Mandate,'* , available at
[126] Profiling is defined under the GDPR as an automated form of processing, carried out on personal data, to evaluate personal aspects about a natural person. Council Regulation 2016/679, *supra* note 13, art. 4(4),
[127] Kaminski, M (2019), *Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability*, available at https://ssrn.com/abstract=3351404
[128] Council Regulation 2016/679, *supra* note 13, arts. 13(2)(f), 14(2)(g),
[129] BIS (2019)
[130] A 'Regulation' unlike a Directive is directly applicable and has consistent effect in all EU member states.
[131] Consumers International (2018) *The State Of Data Protection Rules Around The World A Briefing For Consumer Organizations,* available at https://www.consumersinternational.org/media/155133/gdpr-briefing.pdf
[132] A modern AI/ML methodology is to ingest data, clean label and transform in centralized data hub, explore data, train the data using ML, and then create inference and correlation in the virtual world.
[133] One of the major caveats in using machine learning on real-world data is its continuous demand for high quality annotated training data and, subsequently, the amount of human effort involved in tagging documents manually to create annotations. Privacy Analytics (2019) *Learning at Scale: Anonymizing Unstructured Data using AI/ML*, available at https://bit.ly/2Y5xxfH
[134] Opinion 05/2014 on anonymisation techniques. Technical Report, Article 29 Data Protection Working Party. https://bit.ly/33HLPEu

[135] di Castri, S; Grasser, M and Kulenkampff, A (2018) *Financial Authorities in the Era of Data Abundance: Regtech for Regulators and Suptech Solutions*, available at https://ssrn.com/abstract=3249283
[136] Enhancing model-risk management to address the risks of machine-learning models will require policy decisions on what to include in a model inventory, as well as determining risk appetite, risk tiering, roles and responsibilities, and model life-cycle controls, not to mention the associated model-validation practices. Babel, B; Buehler, K; Pivonka, A *et al* (2019) *Machine learning financial risk management*, available at https://mck.co/2Y9rBmg
[137] A state law in California requires that consumers be informed if they are interacting with an AI-powered chatbot. Other laws may codify the need for fairness and privacy in AI use and decision-making. The law, with certain exceptions, makes it unlawful for any person to use a "bot" to "communicate or interact with another person in California online with the intent to mislead the other person about its artificial identity for the purpose of knowingly deceiving the person about the content of the communication in order to incentivize a purchase or sale of goods or services in a commercial transaction or to influence a vote in an election." See Natlawreview (2019) *California Social Media Bot Disclosure Law Coming Soon*, available at https://bit.ly/2DzsI52
[138] In cryptography, a zero-knowledge proof protocol is a method by which one party (the 'prover') can prove to another party (the verifier) that they know a value x, without conveying any information apart from the fact that they know the value x. It is a core technology in DLT. See Hackernoon (2018) *Explain Like I'm 5: Zero Knowledge Proof (Halloween Edition),* available at https://bit.ly/2DDeq3i
[139] BIS (2019)
[140] Creehan, S and Li, C *(2018)* Asia's Open Banking Push, available at https://bit.ly/2Y76zEs
[141] Consumer Finance (2017) *CFPB Outlines Principles For Consumer-Authorized Financial Data Sharing and Aggregation,* available at https://bit.ly/34EXvJv
[142] *ibid*
[143] *ibid*
[144] As outlined by Cap Gemini (2019) *World Fintech Report 2019*, available at https://bit.ly/2rFGO2e
[145] *ibid*
[146] OAuth 2.0 is an authorization protocol that gives an API client limited access to user data on a web server. GitHub, Google, and Facebook APIs notably use it, and relies on authentication scenarios called *flows*, which allow the resource owner (user) to share the protected content from the resource server without sharing their credentials. For that purpose, an OAuth 2.0 server issues access tokens that the client applications can use to access protected resources on behalf of the resource owner. See further https://oauth.net/2/
[147] BIS (2019)
[148] Cap Gemini (2019) *World Fintech Report 2019*, available at https://bit.ly/2rFGO2e
[149] BIS (2019)
[150] *ibid*
[151] *ibid*
[152] Courts granted a preliminary injunction barring the professional networking platform LinkedIn from blocking hiQ, a data analytics company, from accessing and scraping publicly available LinkedIn member profiles to create competing business analytic products, based on the fact that when a computer network generally permits public access to its data, a user's accessing that publicly available data will not constitute access "without authorization" under the US Computer Fraud and Abuse Act. New Media and Technology Law Blog (2019) *In Blockbuster Ruling, Ninth Circuit Affirms hiQ Injunction -- CFAA Claim Likely Not Available for Scraping Publicly Available Website Data*, available at https://bit.ly/2R6jiG8
[153] Newtech.law (2019) *Can a user's account be accessed through screen scraping?*, available at https://bit.ly/2sy2w8J
[154] Financial IT (2019) *PSD2 Myth Debunking: Screen Scraping will be Forbidden in September 2019,* available at https://bit.ly/2DzG4y6 Newtech.law (2019) *Can a user's account be accessed through screen scraping?,* available at https://bit.ly/2sy2w8J
[155] Financial IT (2019) PSD2 Myth Debunking: Screen Scraping will be Forbidden in September 2019, available at https://bit.ly/2DzG4y6
[156] *ibid*
[157] BIS (2019)
[158] IIF (2018) *Liability And Consumer Protection In Open Banking*, available at https://bit.ly/2L8FLyH
[159] Finextra (2019) *What does open banking mean for cyber security?*, available at https://bit.ly/34EV61I

[160] PwC Canada (2019) *Putting security and privacy at the heart of open banking*, available at https://pwc.to/2DDTtVP

[161] UNCTAD (2017) *Data protection regulations and international data flows: Implications for trade and development*, available at https://bit.ly/2OAdnYk

[162] This could be achieved by entities joining a data protection regime such as the EU-US Safe Harbor Framework/ Privacy Shield, the APEC Cross-Border Privacy Rules system (CBPRs) or a large range of privacy trustmarks schemes), but limiting the scope of their membership to particular activities. *ibid*

[163] Rocher, L; Hendrickx, J, & de Montjoye. Y (2019) *Estimating the success of re-identifications in incomplete datasets using generative models, available at* https://www.nature.com/articles/s41467-019-10933-3

[164] See also Narayanan, A. & Felten, E. W (2014) *No silver bullet: de-identification still doesn't work, available at* https://bit.ly/37VvvTU

[165] Hintze, K & El Emam, K (2018) Comparing the benefits of pseudonymisation and anonymisation under the GDPR, available at https://bit.ly/37WuOtR