



**FIGI** ▶

FINANCIAL INCLUSION  
GLOBAL INITIATIVE



SECURITY, INFRASTRUCTURE AND TRUST WORKING GROUP

# Use of telecommunications data for digital financial inclusion

REPORT OF THE TRUST WORKSTREAM





SECURITY, INFRASTRUCTURE AND TRUST WORKING GROUP

# Use of telecommunications data for digital financial inclusion



## DISCLAIMER

The Financial Inclusion Global Initiative (FIGI) is a three-year program implemented in partnership by the World Bank Group (WBG), the Committee on Payments and Market Infrastructures (CPMI), and the International Telecommunication Union (ITU) funded by the Bill & Melinda Gates Foundation (BMGF) to support and accelerate the implementation of country-led reform actions to meet national financial inclusion targets, and ultimately the global 'Universal Financial Access 2020' goal. FIGI funds national implementations in three countries-China, Egypt and Mexico; supports working groups to tackle three sets of outstanding challenges for reaching universal financial access: (1) the Electronic Payment Acceptance Working Group (led by the WBG), (2) The Digital ID for Financial Services Working Group (led by the WBG), and (3) The Security, Infrastructure and Trust Working Group (led by the ITU); and hosts three annual symposia to gather national authorities, the private sector, and the engaged public on relevant topics and to share emerging insights from the working groups and country programs.

This report is a product of the FIGI Security, Infrastructure and Trust Working Group, led by the International Telecommunication Union.

The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of the Financial Inclusion Global Initiative partners including the Committee on Payments and Market Infrastructures, the Bill & Melinda Gates Foundation, the International Telecommunication Union, or the World Bank (including its Board of Executive Directors or the governments they represent). The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by ITU in preference to others of a similar nature that are not mentioned. Errors and omissions excepted; the names of proprietary products are distinguished by initial capital letters. The FIGI partners do not guarantee the accuracy of the data included in this work. The boundaries, colours, denominations, and other information shown on any map in this work do not imply any judgment on the part of the FIGI partners concerning the legal status of any country, territory, city or area or of its authorities or the endorsement or acceptance of such boundaries..

© ITU 2021

Some rights reserved. This work is licensed to the public through a Creative Commons Attribution-Non-Commercial-Share Alike 3.0 IGO license (CC BY-NC-SA 3.0 IGO). Under the terms of this licence, you may copy, redistribute and adapt the work for non-commercial purposes, provided the work is appropriately cited. In any use of this work, there should be no suggestion that ITU or other FIGI partners endorse any specific organization, products or services. The unauthorized use of the ITU and other FIGI partners' names or logos is not permitted. If you adapt the work, then you must license your work under the same or equivalent Creative Commons licence. If you create a translation of this work, you should add the following disclaimer along with the suggested citation: "This translation was not created by the International Telecommunication Union (ITU). ITU is not responsible for the content or accuracy of this translation. The original English edition shall be the binding and authentic edition". For more information, please visit <https://creativecommons.org/licenses/by-nc-sa/3.0/igo/>

## About this report

This report was written by Rory Macmillan, Partner, Macmillan Keck Attorneys & Solicitors, and Scott Garvey, *of counsel* with the same firm. Vijay Mauree, ITU provided overall guidance for this report. The report was reviewed by the members of the Security Infrastructure and Trust work. The author is thankful to the following people who provided additional comments to the report: Narayan Jaesingh and Jonathan Hakim.

If you would like to provide any additional information, please contact Vijay Mauree at [tsbfigisit@itu.int](mailto:tsbfigisit@itu.int)



# Contents

<b>About this report</b> .....	<b>3</b>
<b>List of Figures</b> .....	<b>6</b>
<b>Acronyms</b> .....	<b>7</b>
<b>1 Introduction</b> .....	<b>8</b>
<b>2 The opportunity of DFS</b> .....	<b>8</b>
<b>3 Telecommunications data</b> .....	<b>9</b>
3.1 Telecommunications usage related data .....	9
3.2 Customer relationship management data.....	10
<b>4 Use of telecommunications data in DFS</b> .....	<b>11</b>
4.1 Customer engagement.....	12
4.2 Credit scoring .....	12
4.3 Risk and asset management more broadly .....	14
4.4 Fraud prevention.....	15
4.5 Identification and AML/CFT .....	15
<b>5 Models for sharing of telecommunications data</b> .....	<b>15</b>
5.1 Data sharing partnership model .....	15
5.2 Data sharing proprietary product model .....	17
5.3 Data interrogation analytics without sharing model .....	17
<b>6 Enabling data to be used to its potential</b> .....	<b>18</b>
6.1 Limitations and incentives.....	18
6.2 The potential of standards .....	19
6.3 A wider vision for use of telecommunications data.....	19
6.4 Regulating to support responsible use of telecommunications data .....	20

## LIST OF FIGURES

Figure 1 – Data Types in a Typical Telecom Provider, Source: Hitachi Vantara and Ravi Kalakota, Transform Telecom: A Data-Driven Strategy for Digital Transformation, June 2019 .....	10
Figure 2 – Development of credit scores with telco data using Machine Learning and agile methodology in Brazil. Source: Luciano Diettrich, Fábio de Souza, and André Guerreiro; Claro Brazil, Paper 4831–2020 .....	13

## Acronyms

DFS	Digital Financial Services
CDR	Call Data Records
UDR	Usage Data Records
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
CPE	Customer Premise Equipment
MFA	Multi-Factor Authentication
MNO	Mobile Network Operator
MSISDN	Mobile Station International Subscriber Directory Number
SMS	Short Messaging Service
USSD	Unstructured Supplementary Service Data

# Use of telecommunications data for digital financial inclusion

## 1 INTRODUCTION

This short paper reviews the use of telecommunications data in digital financial services (DFS).

These data can enable a telecommunications operator and others that have access to such data to assess the regularity and scale of a customer's cash flow, stability of his or her financial condition and nature of his or her social network. Such data can be used to profile the customer for the purpose of targeting services appropriate to his or her needs and capabilities, and offering risk-related services such as credit and insurance.

This offers an opportunity to extend financial services to customers who do not have any other signif-

icant, or even any, transaction record with a bank or other financial institution. If telecommunications can bridge the physical distance between those who live far from financial service providers, the data generated may bridge an information gap between provider and customer.

This paper describes the types of telecommunications data to which telecommunications operators have access, how such data are used in DFS, and how better use of telecommunications data could be facilitated, including through the adoption of standards relating to such data.

## 2 THE OPPORTUNITY OF DFS

Access to basic financial services is a vital part of development, providing the means to manage and protect wealth, invest in the future, and surmount crises. In less-developed economies, large proportions of the population have not had reliable access to such services. Their banks have often been concentrated in urban areas, catering to higher-income individual and business customers. As a result, many low-income people, and people and small businesses in rural areas, have not had bank deposit accounts. As deposit accounts often serve as gateway products to access formal financial services

such as business finance or consumer credit, such customers have been effectively shut out of access to such services.

Banking has been transformed by becoming available online. However, many such services typically require a device capable of using an app or browser rather than a feature phone, and access to the internet. Large amounts of the population remain unable to afford such a device, lack such connectivity, or may struggle with the necessary level of technological or financial literacy.

Financial service providers depend on information to market services to the needs of customers, to assess customers' ability to take on the responsibility of a financial service, to guard against fraudulent transactions, as well as to comply with anti-money laundering (AML) and counter financing of terrorism (CFT) rules. The lack of accounts means a lack of transaction history, and thus often little or no information about a person useful for such purposes.

Widespread adoption of mobile technologies has seen rapid rise in the use of mobile money, especially in sub-Saharan Africa. Mobile money provides many of the functions a deposit account has traditionally provided – and more in terms of easy transfers among peers. Peer-to-peer transfers can provide access to capital for small investments or funding in emergencies such as loss of a job, sickness or death of a family earner. However, mobile money does not provide the same access to credit-based financial services that traditional banking provides.

Individuals and businesses without access to traditional financial services present substantially higher risk for financial services firms than customers that do have such access. They have not established a credit or other financial history with such providers or a credit reference bureau and therefore present greater credit risk and higher underwriting costs.

Mobile telephone penetration is universally higher, and often multiple times higher, than the penetration of bank accounts. Many individuals can thus access a broad range of digital services, including traditional telecommunications services such as voice and data, over-the-top (OTT) services and mobile money services. The data generated by this digital footprint

can be useful to understand customers better. As a result, it can reduce risk and improve availability and accessibility to a range of financial services, particularly credit and insurance.

#### **CDR DATA Consumption features**

##### *By time window and direction*

- Daily call (SMS) events
- Daily duration of call events
- Daily time between consecutive call (SMS) events
- Daily time between consecutive events (either call or SMS)

##### *Global*

- Communications time entropy
- Communications entropy

#### **Social network features**

- Number of unique call (SMS) correspondents
- Call (SMS) delta degrees
- Number of reciprocated call (SMS) events
- Fraction of reciprocated call (SMS) events
- Median of time between reciprocated call (SMS) events

#### **Mobility features**

- Radius of gyration
- Distance traveled
- Popular antennas
- Popular antennas entropy

Source: Pedro, J. S., Proserpio, D., & Oliver, N. (2015). MobiScore: Towards Universal Credit Scoring from Mobile Phone Data. In *User Modeling, Adaptation and Personalization* (pp. 195–207). Springer, Cham. [https://doi.org/10.1007/978-3-319-20267-9\\_16](https://doi.org/10.1007/978-3-319-20267-9_16)

### **3 TELECOMMUNICATIONS DATA**

Telecommunications data includes data held by telecommunications operators about their customers, their accounts, and use of telecommunications services. As illustrated below, telecommunications operators hold a wide variety of types of data.

#### **3.1 Telecommunications usage related data**

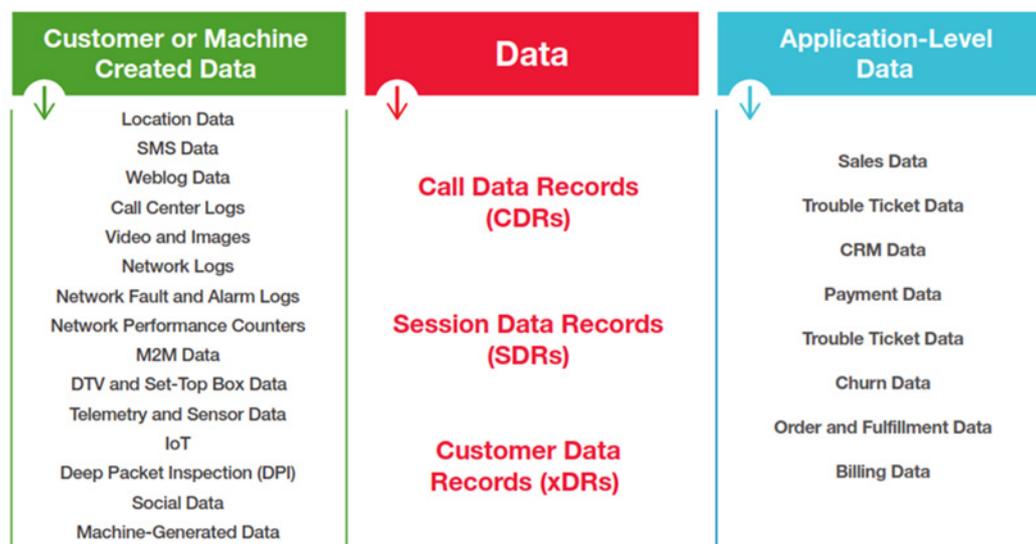
##### **3.1.1 Service usage data**

Telecommunications companies have used call detail records (**CDR**) since before the digital revolution to manage interconnection and roaming domestically and internationally. CDRs track voice call time, date, duration, and initiating and receiving numbers.

Modern telephone CDRs also collect additional data generated by mobile services, including mobile station international subscriber directory number (MSISDN) and location registers, which show roaming locations of mobile users.

CDR data are used not only for billing purposes, but also for telephone accounting and analysis, network management, and fraud detection. For example, CDRs showing call attempts can be used to analyse quality of service and unusual CDR records can be used to detect fraud and plan future capacity requirements. CDR data are kept based on standardized requirements established by the ITU and other regional standards bodies.

Figure 1 - Data Types in a Typical Telecom Provider, Source: Hitachi Vantara and Ravi Kalakota, Transform Telecom: A Data-Driven Strategy for Digital Transformation, June 2019



Operators now use next generation networks that are used for data services. While the term CDR is still used, broader network records are sometimes called generic data records (**XDR**) or usage data records (**UDR**) to denote the much broader range of data that is now routinely maintained. These records include non-voice information such as SMS text messaging data, internet usage and mobility management records.

### 3.1.2 Location data

Even when a mobile phone is not being used to make a voice call or access data, whenever it is switched on it will perform an International Mobile Subscriber Identity (**IMSI**) 'attach procedure' to register on the appropriate mobile network and to verify subscription status. Mobile devices will continue to communicate periodically with mobile networks so long as they are switched on and as they move from place to place, continually updating network registration and subscription status.

### 3.1.3 Network and applications data

Network equipment employed by operators also generate data, including network logs, network fault and alarm logs, network performance counters and deep packet inspection. Communications applications may also generate data. For example, call answering services generate voice recordings and video call services can generate video recordings. SMS records text, and MMS records text, imag-

es and short videos. Equipment used for back-office services also generate data, such as call centre logs.

### 3.1.4 Machine generated data

Telecommunications operators have access to a substantial amount of machine-generated data, and this will only increase as the internet of things (**IoT**) expands and smart devices proliferate. Data is generated by customer access devices, by network equipment and by communications applications. Mobile handsets, IoT devices, DTV set-top boxes, modems, routers and other customer premise equipment (**CPE**) all generate data. This includes device location through communication with cell towers and through telemetry data, and potentially other data generated by sensors.

### 3.1.5 Online activity

Telecommunications operators also have access to data from their customers' online activities through browsers and other applications on their devices. This opens wide opportunities for advertising networks, for example, but is typically restricted by privacy regulations applicable to telecommunications operators.

### 3.2 Customer relationship management data

Operators gather and maintain information for management of the relationship with the customer for commercial and regulatory reasons.

### 3.2.1 Subscriber identification data

At the outset, operators typically collect and hold information about the subscriber, particularly where they are required to do so by know-your-customer (KYC) regulation. Subscriber information will include a subscriber's full name, address, nationality, telephone number, and possibly an email address, a government-issued identification number (such as a passport or National ID number), and biometric information such as a picture, fingerprint, or copy of a government-issued photo ID.

### 3.2.2 Order and billing data

Just as network equipment generates data, applications used to manage telecommunications operators' business operations generate data. This includes sales data, payments data, trouble ticket data, churn data, order and fulfilment data, and billing data. For post-paid accounts, operators have records of customer payment information and potentially other data such as bank or other transaction account information. For pre-paid accounts, which are particularly relevant for financial inclusion, operators have data as to decisions made by subscribers with respect to

calling and data plans and the frequency and value of top-ups. Subscriber information also includes a purchase history for a variety of additional products and services, such as mobile devices and accessories, mobile apps, and other services.

### 3.2.3 Device data

The telecommunications network will record the brand, model and operating system of the device a customer is using. This offers insight as to a customer's level of consumer spending capability or disposable income. Where linked to identity data, it becomes possible to link devices to the same customer. While cross-device identity data is often obtained using email addresses, social media logins and validated linked accounts from different devices, it can also be collected using data signals such as matching locations, IP addresses, types of browsers, and similarities of the operating systems. Cross-device data enables cross-device tracking, fuller data about the customer, and more targeted messaging for instance for customer engagement purposes (see section 5.1 below).

## 4 USE OF TELECOMMUNICATIONS DATA IN DFS

Telecommunications data is used in multiple ways, sometimes on its own as reviewed here. Where people are already users of digital financial services, they will have begun to establish a history of financial transactions. This may begin with use of mobile money, in which case a significant amount of direct financial behavioural history may be combined with the telecommunications data. Where the customer has used digital credit, they will even have a credit history with the lender in question.

The combination of telecommunications data with such financial data is rich. The more the customer builds a credit history with the lender in question, the more weight will be given to that credit history, while the importance of telecommunications data in analysing creditworthiness will recede.

#### CRM DATA Socioeconomic features

- Age
- Gender
- Estimated customer income
- High risk ZIP code
- Regional area code

#### Product features

- Device brand
- Device operating system
- Device type
- Line type
- Line status
- Line quantity
- Late payments
- Month elapsed since activation

Source: Pedro, J. S., Proserpio, D., & Oliver, N. (2015). MobiScore: Towards Universal Credit Scoring from Mobile Phone Data. In *User Modeling, Adaptation and Personalization* (pp. 195–207). Springer, Cham. [https://doi.org/10.1007/978-3-319-20267-9\\_16](https://doi.org/10.1007/978-3-319-20267-9_16)

Nevertheless, a very large proportion of the World's population is not only unbanked but has not used digital credit or even mobile money and still does not have access to smart phones or use many mobile apps. In addition, new customers with no financial history will continue to grow into adulthood. For these reasons, telecommunications data can be expected to remain a vital means of identifying customers and de-risking loans and so lowering their cost for some time to come.

There are several areas in which telecommunications data are used for digital financial services:

- customer engagement, i.e., attracting them to the service;
- credit scoring, i.e., to assess risk of default;
- asset and risk management;
- prevention of fraudulent transactions; and
- customer identification and anti-money laundering and countering the financing of terrorism (AML/CFT).

#### **4.1 Customer engagement**

Telecommunications data is useful to identify and attract potential customers for digital financial services, even basic ones such as mobile money. Mobile operators seeking to launch and grow a mobile money business need to understand their customers, how to prioritise them and market to them. Data scientists have found a significant relationship between mobile telephone usage and the propensity to use mobile money. Analytics firms develop algorithms and identify generic patterns of behaviour and variables that are predictive for identifying users likely to adopt the service. Customers' CDR and CRM data is then calibrated in relation to those models, enabling the mobile operator to target its advertising and bring customers on board.<sup>1</sup> In Uganda, for example, Cignifi partnered with telecom operator Airtel Uganda to use CDRs and mobile money data to identify potential customers that had not registered for lending services.<sup>2</sup>

Analysis of the telecommunications data enables operators to recognize and understand customers

across devices and the customer journeys. Linking offline CRM data with online cookies and mobile devices enables operators to maintain customer experience. Customers can be identified across multiple devices and screens to follow their customer journey and improve their experience as they interact with the operator's brand, including through targeted advertising and transactions.

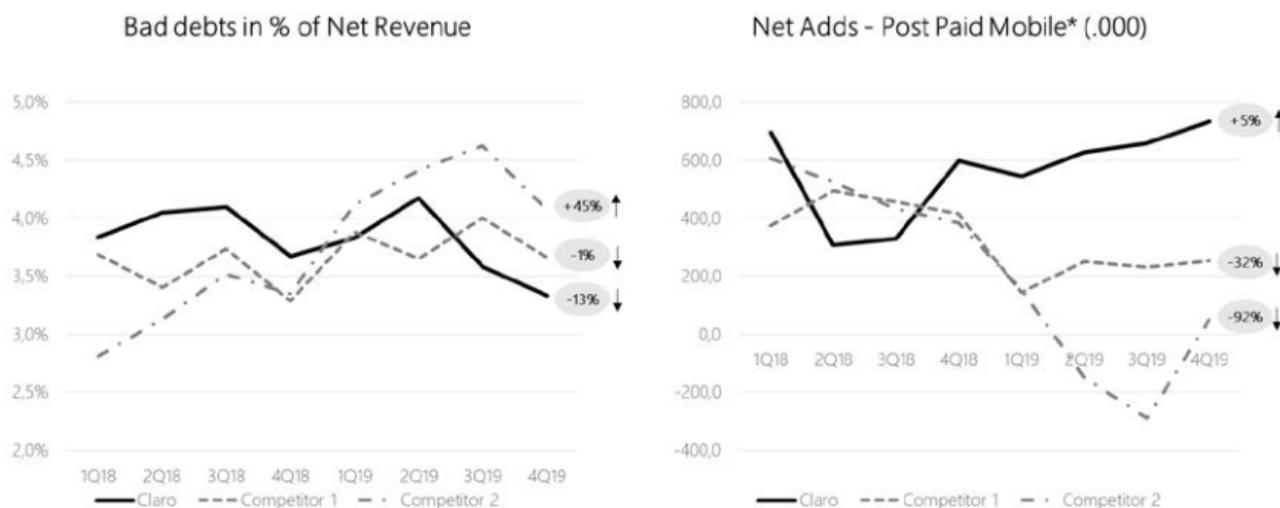
#### **4.2 Credit scoring**

One of the major impediments to further inclusion and deepening is the absence of reliable credit and other information on individuals and enterprises that have not traditionally used banking and insurance services. Telecommunications data can be used to improve access to such services by bridging the information gap between traditional credit information and the data generated by consumers and entrepreneurs. These use their phones not only to make calls and access data, but also to manage their finances, purchase products and services and, increasingly, generate digital data trails.

##### **4.2.1 Post-paid accounts as a form of credit**

Telecommunications data is a useful part of such data trails. To begin with, the switch from a pre-paid to a post-paid account is a transition to a rolling credit account. Telecommunications data is not only useful for assessing a customer's creditworthiness, it is used by the operators themselves for such assessments when considering upgrading a customer from a prepaid account to a post-paid account. Post-paid accounts depend upon the customer to pay his or her debts at the end of each billing period and so the upgrade is effectively a decision to extend credit. Telecommunications operators, such as Claro in Brazil, are able to use machine learning to analyse the telecommunications data they hold on their customers to make better upgrade decisions, and so less customers with defaults on post-paid accounts and so more profitable results.<sup>3</sup> As shown below, the use of telecommunications data reduced Claro's bad debts as a percentage of net revenue by 13%, and enabled a 5% increase in net additions to post-paid mobile accounts.

Figure 2 - Development of credit scores with telco data using Machine Learning and agile methodology in Brazil. Source: Luciano Diettrich, Fábio de Souza, and André Guerreiro; Claro Brazil, Paper 4831-2020



Once upgraded to post-paid account, telecommunications operators (like electric utilities) receive ‘credit like’ payment streams from individuals that could serve as a proxy for data that was more traditionally used in credit scoring, and substantially reduced ‘credit invisibility.’<sup>4</sup> The records of the customer’s usage of telecommunications services constitutes behavioural data rich in potential insights about a person’s wealth and ultimately repayment.<sup>5</sup>

#### 4.2.2 Behavioural insights

For example, a person’s calling behaviour may provide insights into their comparative ability and willingness to repay debt. Initiating larger numbers of calls rather than receiving them, and making of calls of long duration, are used in some data models as supporting a higher credit score.

Customer and billing records offer direct financial data. For pre-paid subscriber accounts, the size and frequency of top-ups and the choice of plan selections illustrate the finances of the user (with some similarities to top-ups of mobile money accounts).<sup>6</sup> For post-paid accounts, subscribers have a credit history based on billing and payment records. Frequent call-backs and use of emergency airtime credit requests enable further compilation of a subscriber’s credit picture. A user who carefully manages his or her prepaid account balance over time to permit smoother usage may be a more responsible borrower. Similarly, where a customer’s service usage patterns follow a monthly cycle, he or she may be more likely to be earning a salary.

Mobility patterns can also show regulatory of employment where travel and location are to a regular location during business hours, supporting stronger credit profiling. Geolocation data about a user, especially when combined with financial data, can also indicate stable housing as well as important socio-economic information such as travel, social and business networks, and other relevant social data such as shopping trends. For example, when a subscriber receives a message asking them to rate a business they just visited, this data is not only used to provide information to other potential patrons, but also to digital service providers.

Telecommunications data also provides valuable social information about a user that is useful for profiling and credit scoring. Family and social networks can be derived from CDRs or calling plans that feature special rates for specific individuals (such as friends, business colleagues or family members). In low-income communities, that social network may be the individual’s financial safety net. Additionally, where others in the community reveal a common pattern of behaviour of financial responsibility and capability, the individual’s profile is strengthened. The strength of an individual’s social connections may be inferred from whether his or her calls to others are returned. Such insights from the data ‘derisk interactions between large firms and the poor, at scale’, enabling ‘new types of formal lending that would not be feasible under historical constraints.’<sup>7</sup>

Machine learning (ML) algorithms are developed using training data. In MobiScore, an AI system which develops credit scores from mobile telephone data, CDRs are used to identify patterns of behaviour that correlate with unreliable financial behaviour. In parallel, credit reports showing actual defaults for the same individuals are used as a ground truth to train the user models.<sup>8</sup> Thus, algorithms segment customers according to a range of behavioural and risk assessment registers.<sup>9</sup> For example, US firm Cignifi worked with Airtel in Uganda to obtain data about the number of calls and text messages made and received per day and phone, web and social network usage, and then analyse that data comparatively using generic models of behavioural patterns.

One study of mobile telephone data on loans to banked and unbanked customers in a middle-income Latin American country showed that such data outperformed traditional credit bureau data: 'Among those with credit histories, if credit were extended to the 50% lowest risk prospects according to the credit bureau the default rate would be 9.7%, whereas it would be only 8.3% based on our scoring using phone records. Moreover, if credit were extended to those without credit histories whose predicted risk of default would place them in the top 50% of risk-prospects for those with credit records, the default rate would be only 6.6%. Our method can identify a group of good credit prospects from among those with no credit history.'<sup>10</sup>

While industry participants remain sceptical of the ability of any CDR-based model outperforming a credit score based on data on past repayment history, there does appear to be a strong opportunity for using such models where such historical repayment data is not available. Safaricom launched M-Shwari in 2012 in Kenya, the first digital credit product that relied on telecommunications data (albeit combined with mobile money usage data) to evaluate risk. Banks, telecommunications operators and insurance companies have also sought to capitalize on telecommunications data to improve delivery of financial services. Telmex, the largest fixed line operator in Mexico and a subsidiary of America Movil, which also owns Mexico's largest MNO, offers small business loans to customers based in part on their phone records.<sup>11</sup> MicroEnsure, a microinsurance firm, has partnered with Telenor Pakistan, the largest MNO in Pakistan, to provide free life insurance for users in

accordance with minimum monthly prepaid airtime purchases.<sup>12</sup>

#### **Using telecommunications data to enable positive credit scoring in Brazil**

The Brazilian positive credit scoring agency Quod is using telecommunications data of the country's telecommunications operators in partnership with US fintech company Cignifi. They are offering credit insights into customers, marketing insights into customers' propensity to certain services, and fraud scores to screen credit applications.

While 97% of Brazilians have mobile phones, 30% do not have bank accounts. The use of telecommunications data can thus enable access to credit for millions of underserved customers who otherwise would not have had such access, and widen the range of products to which individuals and small businesses would otherwise have had access. The partnership's product offerings will include credit insights to complement Quod's positive scores, fraud scores to screen credit applications and on-line transactions, and propensity indicators to enable digital marketing initiatives.

Source: See <https://www.cignifi.com/post/manage-your-blog-from-your-live-site>

#### **4.3 Risk and asset management more broadly**

Telecommunications data can be used not only to reduce risk through credit scoring and profiling, but to provide information about assets that are being financed or insured. For instance, location data is sometimes used to track leased vehicles using starter interrupter devices. Where the customer fails to maintain service on his or her loan, the SID may communicate not only location but may instruct the disablement of the vehicle enabling its recuperation by the lender.

Insurance companies are also already using a variety of IoT data to assess risk. This includes using telemetry data (which can track vehicle location and usage), sensor data from personal fitness devices, smoke detectors, burglar alarms and weather gauges. Most of this IoT data is transmitted over the networks of mobile telecommunications operators and, when used with machine learning, can provide more accurate predictions about insurance claims.

#### 4.4 Fraud prevention

Analysis of location data from mobile telephones indicate that ‘human trajectories show a high degree of temporal and spatial regularity, each individual being characterized by a time-independent characteristic travel distance and a significant probability to return to a few highly frequented locations.’<sup>13</sup> Analysis of mobile phone data of 500,000 Orange customers in Cote d’Ivoire demonstrated that ‘human mobility is highly dependent on historical behaviours and that the maximum predictability is [...] an approachable target for actual prediction accuracy.’<sup>14</sup> That ‘humans follow simple reproducible patterns’ offers potential to identify anomalous behaviour, and so detect potentially fraudulent financial transactions. Data indicating that a person is transacting from an unusual location may be made available to third party financial service providers, such as digital payment service providers and credit card companies, which they may use to block transactions until the individual’s identity is verified.

#### 4.5 Identification and AML/CFT

Digital identity is another area where telecommunications data has great potential to expand financial inclusion. Global efforts to fight terrorism and money laundering have driven “know your customer” (KYC) obligations in several industries, but these KYC obligations do not need to function in industry silos. Internationally, KYC requirements are

mandatory in the banking industry and, increasingly, also for possession of SIM cards. KYC information generally requires a service provider to collect, verify and maintain basic identity information about customers, subject to audit by or sharing with the relevant regulator. The information collected and maintained generally includes verification of the customer’s name, address, nationality, and an official government-issued identification number, such as a passport number. Some jurisdictions also require biometric data such as a photograph or a copy of the customer’s government-issued photo ID.

Many individuals and businesses that require basic financial services do not have traditional bank accounts and therefore no financial institution has collected or verified their KYC information. Additionally, banks have generally instituted KYC systems appropriate to their bricks-and-mortar business model. Mobile operators, on the other hand, have collected and verified KYC information on a much broader segment of the population, including many individuals using mobile money services that do not have traditional bank accounts. The mobile operators have also developed KYC processes consistent with their mobile, digital business models. These MNO KYC processes include appropriate sharing of KYC data with regulators and other third parties such as specialized KYC verification platforms and attention to privacy, data protection and data localization requirements.

## 5 MODELS FOR SHARING OF TELECOMMUNICATIONS DATA

Although MNOs do directly provide certain fintech services, especially mobile money services, the real value of telecommunications data for increased financial inclusion requires a “partnership” between an MNO and a financial services provider (FSP) whereby the telecommunications data is used to improve delivery of the FSP’s financial services. The partnership between the MNO and FSP does not need to be a formal legal agreement or take a particular form, but there needs to be a mechanism for data held by the MNO or insights from it to be shared in some way with the FSP.

We explore here two broad models for sharing telecommunications data for digital financial services:

- Under a formal “partnership” model, an MNO and a financial institution enter into a strategic partnership and share data for that purpose. This

appears to be the earliest form of sharing for DFS.

- Under a data sharing proprietary product model, the MNO does not share any raw data or enter into a formal partnership with a financial institution. Instead, the MNO collects, processes and packages subscriber data into credit scores that can be sold to lenders and other financial services providers as a separate product. In this model, there is no sharing of raw data. The data is shared in the form of the separate credit scoring product.

#### 5.1 Data sharing partnership model

The first model is a formal partnership where the contractual documents forming the legal arrangements set out the terms and conditions for sharing of telecom data for fintech purposes. One of the earliest

and best known uses of this model is the sharing by Safaricom of data generated by its M-PESA mobile money service with commercial banks seeking to use the data in the provision of credit.<sup>15</sup>

### 5.1.1 Business model

In this model, mobile money services are provided by the telecommunications operator, and banking services are provided by the banks. The data sharing is limited to data required for basic subscriber management (e.g. KYC information) and specific mobile money data that the telecommunications operator provides to the banks that the banks can use to make credit decisions. In addition, both parties have access to their own data generated as a result of the partnership (e.g., money transfers between the mobile money wallets and bank accounts).

### 5.1.2 Sharing of data

Under the partnership model, sharing arrangements are defined by commercial agreements. These contain specific principles and procedures for sharing of data and allocate compliance risks and responsibilities. Services provided pursuant to the agreement are explicitly linked to customers' MSISDN, and the set-up of new banking accounts includes provision of full name, ID, birth date, and SIM card IMSI. Each party is required to notify its customers that personally identifiable information is collected and shared with third parties.

A key component of the partnership is for the banks to offer loans to the telecommunications operator's customers, and doing so requires segregating subscribers into different categories based on perceived credit risk. Because these customers have not generated a traditional credit history, telecommunications data, mobile money data and socio-demographic data is used to segregate subscribers.

A credit scorecard is used to rank customers based on unique parameters maintained by the telecommunications operator on subscriber accounts over a given period, e.g., 6 months. The raw data is collected and maintained by the telecommunications operator and processed through application of statistical methods. The banks develop business rules sitting above the scorecard of raw data to convert the credit scores into lending limits for specific customers. The telecommunications operator also shares the credit score parameters used to allocate credit scores to customers and customer demographic details with the banks, and may (but is not required

to) provide additional information such as data bundle purchase and usage patterns, location data and movement data. The banks can review and recommend changes to the scorecard development methodology in line with observed defaults. In addition to collection of credit data, the telecommunications operator is responsible for collection and maintenance of KYC information, which is shared with the banks.

#### Mobile money subscriber and mobile money data shared with banks

- MSISDN
- Data date
- Network registration date
- Payment time
- Number of days below specified levels of airtime
- Airtime utilization
- Activity days and days since last activity
- Average amount of network top-up
- Blocked post-paid numbers
- Network post-paid limit
- Network town
- Information about emergency airtime loans
- Mobile Money registration date
- Mobile Money deposit number
- Mobile Money deposit and withdrawal amounts

### 5.1.3 Telecommunications data shared

In the data sharing partnership model, raw telecom data is collected by the MNO but the financial services provider sets the business rules used to make actual credit decisions. The actual raw data used and shared is specifically set out in the relevant contracts.

Credit decisions are made by the banks based in part on voice, data and mobile money usage on the operator's telecommunications network. There are two main types of data that are shared:

**Mobile money subscriber data** such as the subscriber's name and any of such subscriber's national identity card number, military card number, diplomatic identity card number, alien card number, passport number or driving licence number, as reflected in the telecommunications operator's records from time to time

**Subscriber credit data** is information relating to the use by the subscriber of the mobile money service and the telecommunications network (see text box on the right).

## 5.2 Data sharing proprietary product model

### 5.2.1 Data interrogation analytics without sharing model

An alternative to the partnership model occurs where the MNO does not share raw data, but instead collects and processes its subscribers' data to generate a stand-alone credit scoring product that can be sold to financial institutions or other third parties. KT Corp., South Korea's largest telecom operator, provides a good example of how this model operates. KT Corp. also owns the first internet-only bank in Korea and is active in the fintech, AI and big data analytics spheres. KT Corp. has developed the "K-Telco Score" and the Credit Rating Delivery Platform (CRDP) as alternative credit rating products using telecom big data analysis.

The K-Telco score uses a variety of telecom data, including CDRs, subscriber data, usage data and application data. KT has also commercialized a credit reporting delivery platform (CRDP) that produces the K-Telco score from raw data. CRDP is updated on a daily basis with telecom data, which is synced, processed through the application of big data analysis and machine learning, and packaged into an updated K-Telco score and profile.

KT markets the K-Telco score and CRDP to telecom operators globally as a way to monetize their data. For example, an MNO operating in sub-Saharan Africa could purchase the CRDP and use the platform to generate credit scores based on their subscriber data. Under the business model, when a prospective customer applies for credit, they will provide permission to use personal data for credit purposes. The creditor will request credit data from a credit bureau, which will in turn seek telecom-related credit information. This will be provided by the MNO in the form of the K-Telco score.

There is less transparency to the proprietary product model, but it appears that in the K-Telco score and CRDP appear to seek to make use of the full range of data available to MNO. This would include telecom data, subscriber data, and potentially other data as well. It also isn't clear whether, in the case of the CRDP, KT would have access to the raw data of the MNO using the CRDP.

The proprietary data sharing model can also make use of new technologies, including "federated learning," to improve the accuracy of artificial intelligence. Federated learning is a form of machine learning where artificial intelligence learns algorithms from distributed datasets. In other words, the data sets themselves are never shared, they are only used to train the artificial intelligence algorithms. In China, Tencent's WeBank has used federated learning artificial intelligence to develop small business credit risk models based on distributed data in individual invoice centers. In this case, the raw invoice data in the individual invoice centers is never actually shared with WeBank, but the data is used to develop the artificial intelligence algorithms used by WeBank in its business. Using this approach enabled WeBank to cut its loan defaults in half.<sup>16</sup>

### 5.3 Data interrogation analytics without sharing model

Under another model, a telecommunications operator may permit a financial institution to interrogate the telecommunications operator's data without supplying that data to the financial institution. This widely used approach involves the financial institution installing its business rule engine within the telecommunications operator's premises, and running analytics on it for various use cases. This results in a computed value or output, such as a credit score or confirmation as to regular locations. The financial institution might aggregate the output with other data about the customer, such as KYC data it has already obtained, or credit bureau data to which it has access. Typically, the telecommunications operator will require the customer to consent to his or her data being accessed in this way.

Telecommunications operators may be remunerated on the basis of a revenue share. The details of such revenue share arrangements may be complex to determine, particularly where the financial institution combines the data output with other data before using it, or where the financial institution uses the data for purposes other than immediate revenue-generating services, such as for digital advertising.

## 6 ENABLING DATA TO BE USED TO ITS POTENTIAL

### 6.1 Limitations and incentives

Traditionally, telecommunications operators collected and maintained data necessary to fulfil billing and network management functions necessary for their core business of providing telecommunications services. Today, the costs of collecting and maintaining data have decreased and the potential value of monetizing telecommunications data at scale has increased, creating incentives to use the data created by telecommunications services.

Furthermore, operators can design applications and services specifically designed to facilitate collection of information beyond the purpose of improving delivery of telecommunications services. This potential will only expand with proliferation of machine-to-machine communications (**M2M**). Collection of data may be used to reduce the price of traditional connectivity services, as consumers can 'pay' a portion of the cost of their service by providing valuable data that the operator can monetize in other ways.

There are thus increasing profitable opportunities and so incentives grow to collect, maintain and analyse greater quantities of data. Nevertheless, for a combination of reasons, many telecommunications operators are still not using customer data or making it available to third parties to the extent of its potential.

#### 6.1.1 Declining relative value of telecommunications data

The rising penetration of smartphones, use of mobile data services and mobile apps is creating a vast amount of personal data that offers insights into individual behaviour that far exceeds that offered by telecommunications data described in section 3. DFS providers such as Tala, Branch, Neon, Julo, Grab, Alipay and Konfio are able to use this data powerfully for DFS. The competitive advantage of that data, the increasing prevalence of data-only 4G and 5G networks and the declining role of voice services, means that the relative value of telecommunications data diminishes against the competitive advantage of app usage data. The incentives to use it, then, are lower than might otherwise have been the case.

#### 6.1.2 Risk-averse approaches to regulation

To begin with, most telecommunications operators are subject to confidentiality laws, regulations and licence provisions restricting sharing of data relating to their customers. The customer data that is

subject to such obligations may be defined vaguely or precisely<sup>17</sup> but would typically cover much of the kind of data that would be useful for DFS.

In some cases, such data may be disclosed to other parties where the customer has provided consent to do so. However, this relies on establishing a process that informs the customer with sufficient detail and accuracy what the data will be used for and some form of unambiguous affirmative agreement to use of the data. The combination of cost, time and uncertainty impedes the easy gathering of such data.

Furthermore, while consent is often a lawful basis for processing of personal data under many data protection laws, it has not yet been clearly introduced in many countries' telecommunications laws, regulations and licences as an exception to operators' confidentiality obligations under the telecommunications regulatory framework.

Where there is not a clear and firm legal basis for a telecommunications operator to provide third parties access to telecommunications data, the incentives may not be strong enough for them to take the regulatory risk of doing so. In particular, the revenues that a telecommunications operator can generate from the data may still be only a fraction of the core business of telecommunications services. This leaves those within the organization arguing to exploit the data with a voice that may not be heard amidst the challenges of managing a telecommunications operation.

#### 6.1.3 New regulatory models

Some new regulatory models are emerging that may undermine the immediate value to telecommunications operators of sharing their data. For example, India has introduced a regulated service provider concept of account aggregators. These are licensed entities that obtain the customers' consent to access data about them from various sources. The policy rationale is to ensure that data will be put to economically productive use for the benefit of data subjects while preserving individuals' control of data about them. Where entities hold data about individuals, a regulatory requirement to make that available to other entities could undermine the ability of the former to use the data in a manner that gives them an advantage.

Telecommunications data may thus become treated as a social good rather than a private good. The incentives to generate value may shift from telecommunications operators holding the data to others

having access to it. While this may be a limitation for telecommunications operators use of data they hold, it will be an opportunity for others.

#### **6.1.4 Competitive strategy**

Where operators have recognised the value of the data, they may be competing with others in the services they seek to use it for, such as banks and fintech companies. Telecommunications operators may seek to exploit the specific advantages they enjoy from their privileged control over telecommunications data. Even if they do not do so immediately, they may prefer not to provide access to the data but preserve later opportunities. Mobile operators are not necessarily easy partners, preferring to retain tight control over the customer relationship.

#### **6.1.5 Organisation of data**

Even where telecommunications operators do wish to make full use of data they hold, they may not have organised the data in a manner to make it readily usable for profitable purposes, whether by themselves or third parties. It may be held in different forms, some structured and some unstructured, neither combined nor cross-referenced. In particular, various different data sets about the same customers may not have a common unique identifier (such as a national ID number) enabling insights to be drawn from combining such data.

For all of these reasons, telecommunications data remains under-utilised, and it often takes third party expertise and a partner to prompt the operator to take steps to use the data or make it available to others.

### **6.2 The potential of standards**

Useful work has been carried out to develop standards and ethical guidelines for use of AI in digital financial services.<sup>18</sup> Despite extensive excitement about the potential of data, however, a gap remains with respect to access to data in the manner and scale that uses it to its potential value.

There are various reasons to think that the development of internationally recognised standards could help. Standards could be prepared identifying what telecommunications data should be made available, and setting out how it should be collected and organised by telecommunications companies.

To begin with, not all data are as useful as others. Standards might help form a consensus focused on data that really have value, such as certain call patterns or billing data, and to prioritise organising these so they can be available in a common format.

Standards also might identify certain data that are potentially very sensitive, such as location data, and build in measures to protect such data such as criteria for access to them, controls on the purposes for which they may be used, and how such data should be deleted. Anonymisation and other privacy enhancing technologies might also be agreed upon for standard use.

Standards could also provide for the use of unique identifiers (e.g., national ID numbers) across data sets to allow for richer layers of data to be used, but coupled with careful measures to mitigate risks to privacy from combining personally identifiable information.

Some basic parameters for explaining to consumers how data has been used could be included in standards.<sup>19</sup> Standards could also set out principles or mechanisms enabling customers to have recourse if data about them is used without a proper legal basis, or is misused, or is incorrect (e.g., due to being out of date or erroneous).

Such standards could also set out the types of organisations that could access such data, and the purposes for which they might legitimately do so. This could include standards for verifying such organisations and their claimed purposes, for authenticating their identities, for controlling which data they may have access to, and measures for limiting their use and access to the purpose for which it is permitted.

Such common standards would serve several purposes of improving the usefulness of data when made available. If implemented widely, this would create an externality of increasing the overall value of all data for social benefit because data are so much more useful when aggregated at large scale. Common standards would also support the legitimacy of using such data. Regulatory authorities could endorse the international standards, or even adopt them into their domestic regulations.

These measures should significantly reduce the regulatory risk that telecommunications operators face. The creation of value and reduction of risk would significantly increase the incentives for telecommunications operators to use and allow third parties to use data they hold.

### **6.3 A wider vision for use of telecommunications data**

The development of standards would also encourage telecommunications operators to see the opportunity of the data they hold at many levels, beginning (where they do not already) to improve profitability of their core telecommunications business. They may

benefit from analysis of their customers using training data and big data algorithms developed across all operators in a country in order. The strong improvement in Brazilian operator Claro's financial results from upgrading prepaid to post-paid accounts mentioned in section 4.3 is only part of the opportunity. In addition, telecommunications operators can use such analytic tools to optimise the customer lifecycle, minimising churn and increasing lifetime customer value, as well as obtaining more deeper and more accurate insights into their customer base.

This greater attention to the value of telecommunications data, combined with standards and in due course regulatory endorsement, can reasonably be expected to lead to greater use and ultimately improvements in a wide range of services.

The generation of revenue from telecommunications data may also lead to a shift in the value of providing telecommunications services from collection of consumer payments for mobile or data services to the collection of data that can be monetized by telecommunications companies in other ways. If the cost of providing telecommunications services declines relative to the value of data being collected, telecommunications operators may adjust their commercial practices to encourage greater uptake of mobile services that generate the data and use of mobile handsets that are capable of generating valuable data. This could result in more smartphone subsidies, and contribute to narrowing the digital divide.

## **6.4 Regulating to support responsible use of telecommunications data**

### **6.4.1 Data for social value**

In addition to endorsing standards discussed in section 8.2, regulatory authorities in telecommunications and financial sectors might be encouraged to consider other ways to exploit telecommunications data for greater social value in DFS. For instance, Brazil introduced in 2019 a requirement for telecommunications operators and other utilities to report positive credit data (i.e., not merely negative data recording defaults but positive data recording successful payment of debts) for credit reference purposes. The availability of this data at scale enables firms to achieve considerably more precise assessments of customers' likelihood to repay debt. Because it is industry wide, it produces a very large volume of data about the population, lending itself then to big data analytics.

### **6.4.2 Consumer protection and privacy**

The more that telecommunications data is used for purposes beyond managing the provision and marketing of telecommunications services, the more important it will be to ensure it is subject to safeguards. Consumer protection and privacy rights must be respected, of course, such as ensuring that consumers are aware what data are collected about them and their activities, and for what purpose they are collected, that only such data as are required for such purposes are collected, and that the use of the data is legitimate (whether because authorised by law, the consumer's consent or otherwise).

The G20 High-Level Policy Guidelines on Digital Financial Inclusion for Youth, Women and SMEs includes Featured Policy Option No. 3.7: "Ensure the responsible use of alternative data, consistent with applicable laws and good practices related to consumer protection, and remain vigilant to potential financial stability risks."<sup>20</sup>

However, ensuring informed consent to use of any data presents a challenge for financial inclusion. On one hand, informed consent is widely used as a touchstone to protect individual data privacy rights. Yet in the context of a mobile application, the consumer's transaction costs associated with obtaining and providing informed consent may result in a pro forma click.

#### **Using telecom data for DFS in Brazil**

Data analytics firm Cignifi has partnered with four mobile operators. They are producing scores for fraud purposes. Using a collaborative model that includes all four mobile operators in the market provides scale and increases the reliability of the data, increasing its value to lenders. The credit bureau hosts the data and offers low cost distribution to major consumers of the data such as banks and insurance companies. Brazilian data privacy rules permit use of personal data for fraud and credit without requiring customer consent. Data is interrogated using an "enquiry only model" layer of software to query the databases for Yes/No answers without sharing the raw data itself, e.g., to verify whether a cellphone number matches a national ID number. This allows the creation of value from extraction of sensitive data from its source while preventing disclosure of unnecessary personal data. For example, it enables the analytics to produce a localization score validating the home or work address of a customer using his or her geolocation pattern as a mobile phone user. This is then used for onboarding the customer for ecommerce and financial products.

Many of these risks are aggravated by the context in which telecom data is used. Subscribers may be illiterate or unfamiliar with the services they need. They are also likely to be using old technology and low-cost data services that cannot transmit the quality and quantity of information consumers need, because most subscribers using mobile money and fintech services are not using smartphones and the relevant technology does not have the ability to transmit large data files such as disclosure and consent forms. For example, a study in Tanzania revealed that only 41% of customers were informed of terms and conditions when signing up for MNO-led insurance product.<sup>21</sup> Standards, discussed in section 6.2, may be helpful in this area, failing which it may be necessary to introduce greater regulation to ensure proper information is provided and consent obtained.

The weakness of consent as a justification for using personal data raises the question whether it is worth requiring consent at all in some cases. The stronger the social benefit of using the data, the weaker the rationale is for protecting the consumer who is in reality not meaningfully protected anyway. It will be worth exploring whether the emphasis should rather be placed on building privacy protections into data access in a manner that minimizes disclosure of personal data while maximizing its utility.

The project in Brazil described in the accompanying box suggests that it is indeed possible to use personal data without the customer's consent while still minimizing the privacy risk to the individual. In that project, the credit bureaus remain key, and thus data governance placed within a centralized, regulated institutional context. Greater flexibility to make data available while securing privacy might be achieved also through using publicly owned data collaboratives relying on robust frameworks setting out rights and responsibilities for access to the data and the processing that can be applied to it.

#### **6.4.3 Financial stability and alternative data**

The G20 Guidelines recognize that fintech companies providing credit and insurance services have developed credit assessment models that use alternative data, in particular mobile phone data, to evaluate risks. The G20 noted concerns about explicability and bias, and also that these data sets have been used only during the period of positive economic growth since 2008 and therefore have not been stress-tested against potential negative economic periods, which could result in financial stability risks. However, systemic risks associated with alternative data may

increase if alternative data becomes fundamental to credit decisions. Because alternative data is intended to provide greater inclusion for low-income and historically disadvantaged populations, the negative consequences of systemic shocks resulting from alternative data may disproportionately harm those who can least afford it.

#### **6.4.4 Cyber security**

As telecommunications data becomes more sought after, cyber security will become more important, both for telecommunications operators and their partners who are permitted to access their systems. This may not be a matter only of preventing fraud from theft or leakage of data, but also of manipulation of telecommunications data to obtain desirable automated decisions made by digital financial service providers in reliance on such data.

#### **6.4.5 Data localisation**

Data localisation requirements are increasingly being legislated and enforced, often with unforeseen consequences. This and other restrictions on transporting data outside a jurisdiction may also present impediments to use of telecommunications data for greater financial inclusion. This is likely so particularly where provider groups operate in multiple jurisdictions. The combined telecommunications data of a telecommunications group active in several countries in a region is likely considerably more useful than keeping the data in national silos. Likewise, making telecommunications data available to digital financial service providers and their technical partners at scale across multiple countries allows them to draw more precise insights about their potential customers. The importance of allowing cross-border transfer of data for digital financial services, including for inclusion, is increasingly widely recognised.<sup>22</sup>

#### **6.4.6 Identification systems**

The development and adoption of standards providing for use of unique identifiers across organisations and data sets was mentioned in section 8.2. This might be achieved by, or coordinated with, the adoption of national identification systems and ID numbers where these are being developed.<sup>23</sup> Use of unique identifiers for telecommunications data on a national basis may also reduce risks of error to the consumer of digital financial services. The use of other identifiers such as MSISDN (mobile telephone numbers) carry high risks, particularly where numbers may be transferred to other people.<sup>24</sup>

#### **6.4.7 Competition**

Lastly, if it becomes apparent that control over data by a dominant operator confers on it dominance in adjacent markets (e.g., provision of digital financial services), regulating to avoid one firm earning super-normal profits at the expense of consumers may be

necessary. Intervention of competition authorities may also be necessary if there is evidence that data is being held and used in ways that harm competition. These are complex areas of economic regulation that are not the subject of this paper but may be appropriate to consider in some markets.

## Endnotes

- <sup>1</sup> <https://www.cignifi.com/>
- <sup>2</sup> Cignifi and the IFC to Grow Usage of Mobile Money in Uganda. Cambridge, MA: Cignifi. <https://www.prnewswire.com/news-releases/ifc-partners-with-cignifi-to-grow-usage-of-mobile-money-in-uganda-300030886.html>
- <sup>3</sup> Luciano Diettrich, Fábio de Souza, and André Guerreiro; Claro Brazil, Paper 4831 - 2020, *Development of credit scores with telco data using Machine Learning and agile methodology in Brazil*, available at <https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2020/4831-2020.pdf>.
- <sup>4</sup> Turner M, Lee A, Schnare A, et al. (2006) Give Credit Where Credit Is Due: Increasing Access to Affordable Mainstream Credit Using Alternative Data. Washington, DC: Political and Economic Research Council/The Brookings Institution.
- <sup>5</sup> Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, 350(6264), 1073-1076.
- <sup>6</sup> Mobile money accounts also show financial transactions, and for many users can provide the same picture of finances as a traditional deposit account, including size and frequency of deposits and transaction history.
- <sup>7</sup> Bjorkergren D and Grissen D (2015) Behaviour Revealed in Mobile Phone Usage Predicts Loan Repayment, available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2611775](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2611775).
- <sup>8</sup> Pedro, J. S., Proserpio, D., & Oliver, N. (2015). MobiScore: Towards Universal Credit Scoring from Mobile Phone Data. In *User Modeling, Adaptation and Personalization* (pp. 195-207). Springer, Cham. [https://doi.org/10.1007/978-3-319-20267-9\\_16](https://doi.org/10.1007/978-3-319-20267-9_16)
- <sup>9</sup> Bjorkergren D and Grissen D (2015) Behaviour Revealed in Mobile Phone Usage Predicts Loan Repayment, available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2611775](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2611775).
- <sup>10</sup> Bjorkergren D and Grissen D (2015) Behaviour Revealed in Mobile Phone Usage Predicts Loan Repayment, available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2611775](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2611775).
- <sup>11</sup> The Future of FinTech: A Paradigm Shift in Small Business Finance 19 (World Economic Forum 2015)
- <sup>12</sup> Innovation in Electronic Payment Adoption: The case of small retailers 25 (World Bank Group June 2016)
- <sup>13</sup> Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *Nature*, 453(7196), 779-782. <https://doi.org/10.1038/nature06958>
- <sup>14</sup> Lu, X., Wetter, E., Bharti, N., Tatem, A. J., & Bengtsson, L. (2013). Approaching the Limit of Predictability in Human Mobility. *Scientific Reports*, 3. <https://doi.org/10.1038/srep02923>
- <sup>15</sup> Safaricom entered into partnership agreements with the Commercial Bank of Africa (CBA) and the Kenya Commercial Bank (KCB) to offer M-PESA customers access to banking services, including savings products and loans. The CBA agreement was entered into in 2013. The KCB agreement was entered into in 2015.
- <sup>16</sup> *Tencent's WeBank applying "federated learning" in A.I.* (DigFin 29 July, 2019) available at <https://www.digfingroup.com/webank-clustar/>.
- <sup>17</sup> For example, Section 222(h)(1) of the US Communications Act defines customer network proprietary information (CNPI) is: "(A) information that relates to the quantity, technical configuration, type, destination, location, and amount of use of a telecommunications service subscribed to by any customer of a telecommunications carrier, and that is made available to the carrier by the customer solely by virtue of the carrier-customer relationship; and (B) information contained in the bills pertaining to telephone exchange service or telephone toll service received by a customer of a carrier; except that such term does not include subscriber list information."
- <sup>18</sup> E.g., Monetary Authority of Singapore, *Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector*, available at <https://www.mas.gov.sg/-/media/MAS/News%20and%20Publications/Monographs%20and%20Information%20Papers/FEAT%20Principles%20Final.pdf> and *Smart Campaign Digital Credit Standards*, available at <http://www.smartcampaign.org/>.
- <sup>19</sup> A wide range of risks associated with artificial intelligence have been identified and data regulators internationally are taking steps to address them. These include the limited "explainability" of decisions resulting from machine learning, the potential for bias in datasets (including bias with respect to gender or ethnicity) to result in biased decision-making, the problems of securing informed consent, and application of cyber security requirements. Use of telecommunications data for fintech presents certain heightened risks. For example, individuals benefiting from financial inclusion may be illiterate or at the very least unfamiliar with fair practice with respect to financial services. Use of machine learning, especially federated learning, increases the absence of explainability in credit decision-making.
- <sup>20</sup> G20 High-Level Policy Guidelines on Digital Financial Inclusion for Youth, Women and SMEs 26 (G20, 2020)

<sup>21</sup> Mobile Insurance Regulation at 25.

<sup>22</sup> G20 High-Level Policy Guidelines on Digital Financial Inclusion for Youth, Women and SMEs featured policy option 3.4 (p25), to “Improve availability and accuracy of SME information, expand credit information sharing, and enable responsible cross-border data exchanges.”

<sup>23</sup> The form of KYC information may also vary across sectors. These may also unwittingly present barriers to greater inclusion. Digital identification technology, including biometric identification cards is critical (lack of government-issued IDs are a significant barrier to financial inclusion where present, such as Sub-Saharan Africa). Indeed, for traditional bank accounts, a government-issued ID might not be enough, to open an account. Financial institutions will often require customers to provide proof of local residence by showing, for example, a utility bill. These challenges are widely recognized and being addressed through new comprehensive national digital identification legal frameworks.

<sup>24</sup> As noted above, the M-PESA partnership is explicitly linked to MSISDN, which are conventionally known as mobile telephone numbers. However, mobile subscriptions have a high turnover rate, and mobile numbers are recycled by operators after they have been dormant for a prescribed period. However, credit accounts tied to MSISDNs have followed phone numbers instead of subscribers. In Africa, new Safaricom subscribers have received collection texts for collection of unpaid debts by the previous user of their phone number.





International Telecommunication Union  
Place des Nations  
CH-1211 Geneva 20  
Switzerland