DISCUSSION NOTE

Financial Consumer Protection and New Forms of Data Processing Beyond Credit Reporting

NOVEMBER 2018

WORLD BANK GROUP

Ministry of Foreign Affairs of the Netherlands

Finance, Competitiveness & Innovation Global Practice

This work is a product of the staff of The World Bank with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent.

The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

RIGHTS AND PERMISSIONS

CONTENTS

BOXES AND FIGURES

ACKNOWLEDGMENTS

This discussion note is a product of the Financial Inclusion, Infrastructure & Access Unit in the World Bank Group's Finance, Competitiveness & Innovation Global Practice. The note was prepared by Rosamund Clare Grady, Fredesvinda Fatima Montes, and Marco Traversa. Douglas Pearce provided overall guidance.

The team is grateful for peer review comments received by Margaret Miller and David Medine.

The team benefited from discussion and consultation papers, and from presentations and discussions at events and conferences, including the Responsible Finance Forum VIII (April 2017) and the Consultative Group to Assist the Poor and International Telecommunications Union's joint data-privacy and digital-finance workshop, "Two Data-Driven Financial Inclusion Business Models and Implications for Data Privacy" (April 2016). The team gratefully acknowledges useful comments received from members of the International Committee on Credit Reporting (ICCR), the International Financial Consumer Protection Organistion (FinCoNet), and the G20-OECD Task Force on Financial Consumer Protection.

The team thanks Naylor Design, Inc. for design and layout assistance, and Charles Hagner for editorial inputs.

Finally, this report would not be possible without the generous support of the Netherlands' Ministry of Foreign Affairs.

ACRONYMS AND ABBREVIATIONS

APEC	Asia-Pacific Economic Cooperation
ARCO	access, rectification, correction, and opposition
DFS	digital financial services
FCP	financial consumer protection
G20	Group of Twenty
G20 DFI HLPs	G20 High-Level Principles for Digital Financial Inclusion
G20 FCP HLPs	G20 High-Level Principles on Financial Consumer Protection
General Principles	General Principles for Credit Reporting
Good Practices	Good Practices for Financial Consumer Protection
GSMA	GSM Association
ITU	International Telecommunications Union
KYC	Know Your Customer
OECD Guidelines	OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data

1 INTRODUCTION

The objective of this discussion note is to provide an overview of consumer-related benefits and risks arising from the usage of new types of data, beyond traditional credit-reporting frameworks, for the provision of financial services, while also aiming at identifying areas for further research. The note was developed primarily by synthesizing discussions about new forms of data—including big data–related financial consumer protection (FCP) issues—both in the various international forums in which the World Bank Group is represented and in consultation and discussion papers. Special regard has been given to the likely benefits and risks for consumers in developing countries and emerging markets. Consideration has also been given to international FCP standards and good practices.

Technological innovation in the financial sector is a global and rapidly growing phenomenon, with particular relevance for developing economies and emerging markets. The use of new technologies is changing how financial products and services are being designed and delivered, vastly increasing the potential number of users by allowing for access even in remote, rural locations while reducing the cost of services. These innovations are especially relevant for developing economies and emerging markets, encouraging them to embrace digital financial-inclusion strategies, with all their potential for economic growth and poverty reduction.[1]

A key part of this technological innovation is the usage of new types of data and data-processing tools, including big data, in the provision of financial services. While the most talked-about issue is "big data," focusing on data sets that are characterized by exponential growth in volume, variety, and velocity and that are the subject of advanced data analytics, many other types of new data, beyond traditional credit-reporting frameworks, are currently being used and processed for the provision of financial services.

New technologies, combined with the usage and processing of new types of small-, medium-, and large-volume data, can support financial inclusion and bring benefits to financial consumers. New sources of data, as well as new ways of processing such data, are a key contributor to the explosion in the accessibility of convenient and tailored digital financial services (DFS) to served, underserved, and unserved consumers. They are being used to;

- Design and market "consumer-centric" (digital) financial services for the unbanked;
- Create credit scores for consumers without a formal credit history or with limited credit hiostry;
- Meet and facilitate compliance with "Know Your Customer" (KYC) requirements;
- Price financial services to reflect the risk profile of individual consumers; and
- Minimize the risk of fraud.

On the other hand, issues may arise, as there is a great variety of personal information that may be used and processed in this context. First, traditional financial services providers generally have access to conventional forms of highly organized, readily searchable, structured data, including client data—credit history and scores, IDs, demographics, and survey data—as well as transactional data. Second, technological advancements allow the usage of other forms of data that are new in the sense that they either have not (until recent times) been used by financial services providers or are not necessarily related to the use of financial services. They include, for example, social-media data and data about the usage of e-money, air time, online search and shopping habits, utility payments, psychometric data, Internet-based entertainment services, devices connected to the Internet of Things, and

[1] Arjuna Costa, Anamitra Deb, and Michael Kubzansky, "Big Data, Small Credit: The Digital Revolution and Its Impact on Emerging Market Consumers" (Omidyar Network, 2016), 6. See also Tavneet Suri and William Jack, "The Long-Run Poverty and Gender Impacts of Mobile Money," *Science* 354, no. 6317 (December 9, 2016), 1288–1292.

data used to determine insurance-related risks. The power of some of this data—such as information on business cash-flow and sales history, which is available through e-commerce platforms—has created opportunities for new entrants to financial services. Some of these "new types of data" may also be unstructured (such as emails, texts, audio files, digital pictures, videos, and messages). And, finally, public sources of data are also available from courts and bankruptcy records, all forms of media, and electoral rolls, which could also be exploited and taken into consideration in big-data analytics.

Given the lack of a framework, risks for financial consumers are growing in nature and scale as the use of these new types of data expands. The recent increase in the aggregation and analysis of huge volumes of diversely sourced personal information, and the speed with which it is processed, create the risk that individuals will be defined by reference only to data and algorithms, rather than personal information. More specifically, key risks include the following: uninformed and meaningless consumer consent to the use of personal information; illegal discrimination; unfair price segmentation; lack of transparency about the collection, use, and disclosure of personal information; insufficient data security (the greater the volume of data being stored, the greater the risks); and failure to provide effective access and correction and complaints-handling mechanisms. The potential for these risks to cause harm is greater where consumers have low levels of financial capability, as is the case in many developing economies and emerging markets.

A further factor affecting the growth of benefits and risks is the rapidly expanding use of smartphones to deliver DFS. As of December 2016, there were around 2 billion smartphones globally, and this number is expected to rise to 4 billion by 2020, with much of the growth in emerging markets and developing economies.[2] This growth is fueled largely by the decreasing cost of smartphones, which generate valuable data that can facilitate cheaper, more tailored financial services, and many economies are recognizing this potential.[3]

Given the increased usage of new types of small-, medium-, and large-volume data and considering its potential benefits and risks, there has been growing international focus on this phenomenon. The Group of Twenty (G20), the World Bank's Finance and Markets Global Practice, the Bank for International Settlements, the Consultative Group to Assist the Poor, the Better Than Cash Alliance, the U.S. Consumer Financial Protection Bureau, the GSM Association (GSMA), the European Banking Authority, and the International Monetary Fund, among others, have provided guidance and launched workstreams on this topic. For example, the GSMA's Code of Conduct for Mobile Money addresses transparency of data, user choice and control, data minimization, fraud management, and security.[4] The Better Than Cash Alliance's Digital Payments Guidelines also call for clients' digital data to be kept confidential and secure.[5] Importantly, the G20 High-Level Principles for Digital Financial Inclusion (G20 DFI HLPs) state: "Digital technology also enables the generation and analysis of vast amounts of customer and transaction data ('Big Data'), which introduces its own set of benefits and risks that should be managed."[6]

While there is existing guidance on the use of personal data and recognized financial consumer rights in existing frameworks and international standards, there is a need to examine the frameworks and their adequacy in relation to this phenomenon in the financial sector. Relevant issues may include those relating to consumer and data protection, privacy, credit reporting, competition, and discrimination. In addition, sectoral standards and regulations covering aspects of payment systems, credit information, and data analytics might also include provisions on the massive use of data from different sources. Generally, robust data-protection standards cover only some of these issues, and even where such standards are robust, a key aspect to consider is the implementation of existing rules and principles, particularly relating to FCP, data privacy, and credit reporting, in the context of big data. Finally, beyond the application of rules and principles, this phenomenon poses challenges to responsible authorities both within the national context and internationally and requires firm collaboration between them,[7] as even where regulatory frameworks exist, issues are likely to stretch the supervisory capacity of relevant regulators.

[2] See, for example, https://marketrealist.com/2016/12/smartphones-changing-ems-business-landscape

[3] See, for example, recommendations from the various committees of the Government and Reserve Bank of India, which are working on financial inclusion and recommending linking data generated by unique personal identifiers to credit bureaus.

[4] Code of Conduct for Mobile Money Providers, Principle 8 (GSMA, 2015), available at https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2015/10/Code-of-Conduct-for-Mobile-Money-Providers-V2.pdf

[5] Responsible Digital Payments Guidelines, Guideline 7 (Better Than Cash Alliance, 2016), available at https://www.betterthancash.org/tools-research/case-studies/responsible-digital-payments-guidelines

[6] G20 High-Level Principles for Digital Financial Inclusion, Principle 5 (Global Partnership for Financial Inclusion, 2016), available at https://www.gpfi.org/publications/g20-high-level-principles-digital-financial-inclusion.

[7] See, for example, "Joint Committee Discussion Paper on the Use of Big Data by Financial Institutions," (European Securities and Markets Authority, European Banking Authority, and European Insurance and Occupational Pensions Authority, 2016).

2 SCOPE AND DEFINITIONS

In recent years, financial services providers have begun using and processing different types of data to make financial services and products more cost-efficient and tailored to consumers' needs. Given new technological advancements, including the advent of FinTech, and the uptake in electronic transactions and electronic commerce, more and more consumers are generating data that is in turn being processed and used by financial services providers and FinTech. While no single definition of data is used, it can include small, medium, and big data, as well as both unstructured and structured data. It can be data on electronic payments or remittances recipients, as well as social-media data and other types of data, including Internet searches, online shopping, and so forth. The issues presented in this background document cover all types of data and are not limited just to big data, although given its complexity, it deserves special attention.

Given big data's increased relevance, continued usage, and complexity, and considering that definitions vary across sectors, regulatory bodies, and countries, it is important to define *big data* before commencing the analysis.

Definitions typically focus on the three Vs—the volume, velocity, and variety of the collected data and the related advanced processing techniques. The Gartner definition is often quoted in this context: "Big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making and process automation."[8] Governments and regulatory authorities have adopted and used similar definitions. Below is a more detailed explanation of the three Vs:[9]

- **Volume.** Although there are diverging views about the total volume of new data created on the web, an often-cited estimate is that at least 2.5 exabytes are generated every day, and it's predicted that 40 zettabytes will be created by 2020.[10] This shows that an immense amount of data is being produced and accumulated, and that the amount is growing very rapidly.
- **Velocity.** This refers to the rate at which data is generated.
- **Variety.** This term refers to the wide range of data being collected, analyzed, and used. Examples include traditional data from financial transactions as well as data from social-media networks, psychometric testing, air-time usage, mobile phone and email communications, chat sites, online shopping habits and transaction histories, gaming sites, use of Internet-based entertainment services, virtual currency transactions, utility payments, and data from the Internet of Things. Information on shipping and delivery times, consumer reviews, complaints data, mobile money transaction volumes, and other data relevant to small businesses run by individuals is also relevant.

Beyond the three Vs, big data is also characterized by advanced data analytics and related algorithms. Examples include using algorithms to find correlations in a form of machine learning; collecting and analyzing all available data, rather than, for example, sampling the available data randomly; and repurposing data—that is, using data provided for one purpose for another (using social-media data for marketing purposes, for example).[11]

Additionally, the term *big data* covers both structured and unstructured data. *Unstructured data* has been usefully defined as "referring to information that either does not have a pre-defined data model and/or is not organized in a predefined manner."[12] Examples given include emails, text files, audio files, presentations, digital pictures and videos, images and messaging, as well as potentially the underlying sources of metadata.[13] *Structured data*, on the other hand, has been defined as (in summary) "information with a high degree of organization, such that inclusion in a relational database is seamless and readily searchable by simple, straightforward search-engine algorithms or other search operations."[14]

For the purpose of this document, the *consumer* is understood as a person or a micro- and/or small enterprise whose data may be collected, used, and disclosed for personal or business purposes.

[8] See "Big Data" in the Gartner IT Glossary, available at **http://www.gartner.com/it-glossary/big-data.** See also **https://esas-joint-committee.europa.eu/Publications/Discussion%20Paper/jc-2016-86_discussion_paper_big_data.pdf**

[9] See, for example, the definition used in "Joint Committee Discussion Paper."

[10] To put this in context, an exabyte has been estimated as 1 trillion, 600 billion books, or about 3,000 times the content of the Library of Congress. Another statistic to note: By 2025, there could be up to 75.4 billion devices connected to the Internet (up from 15.4 billion in 2015). Louis Columbus, "Roundup of Internet of Things Forecasts and Market Estimates, 2016," Forbes, November 27, 2016, available at **https://www.forbes.com/sites/louiscolumbus/2016/11/27/roundup-of-internet-of-things-forecasts-and-market-estimates-2016/#b65fac4292d5**

[11] "Big Data and Data Protection" (Information Commissioner's Office, 2014), 9.

[12] Michelle Nemschoff, "A Quick Guide to Structured and Unstructured Data," Smart Data Collective, June 28, 2014, available at **http://www.smartdatacollective.com/michelenemschoff/206391/quick-guide-structured-and-unstructured-data.**

[13] Nemschoff, 2014.

[14] "Structured vs. Unstructured Data," BrightPlanet, June 28, 2012, available at **http://www.brightplanet.com/2012/06/structured-vs-unstructured-data/.**

3 CONSUMERS' VIEWS ON SHARING THEIR PERSONAL INFORMATION

While the usage of new and different form of data may bring advantages for financial inclusion and enhance product suitability, recently conducted surveys suggest that consumers have different attitudes toward sharing personal information to receive a better offer. Surveys conducted by the Boston Consulting Group, GSMA, the International Telecommunications Union (ITU), and the Omidyar Network show that, overall, consumers care about their data. Despite this, they are often likely to consent to sharing their data without reading the terms and conditions first, and the extent to which they may be willing to share certain types of information varies significantly.

As consumers are asked more and more often to share their data, it can be inferred that, while they consider certain data private and sensitive, the majority believes that sharing is part of everyday life. A Boston Consulting Group survey conducted in 2013[15] showed that consumers see data related to their phone and financial-services usage as moderately to extremely private (89 percent and 65 percent, respectively), but less than 50 percent of consumers see social-network data and information about purchases as moderately to extremely private. Having this in mind, 78 percent of consumers understand that sharing data is part of everyday life.

A similar survey focused specifically on mobile Internet usage.[16] It showed that while most consumers are concerned about sharing their data, a smaller proportion will check what information is required to be shared and related policies and procedures before installing an application. However, a significant number of consumers would like to have consistent data-protection and privacy-related rules. In fact, according to the survey, over 80 percent of mobile Internet users worldwide had concerns about sharing their personal information when accessing mobile applications and services. This number dropped to 65 percent when only those mobile-application users who check what information an application wants to access were taken into consideration. Similarly, 81 percent of mobile users think it is important to be informed and to have the option of agreeing each time their personal information is shared with third parties. Nevertheless, most mobile-application users with privacy concerns (52 percent) would still use the application. Lastly, 60 percent of mobile users want a consistent set of rules to apply to any company accessing their location, regardless of how they obtain this information.[17]

Another survey, conducted in 2016 by the Omidyar Network, showed that more and more consumers use their phones to make financial transactions and consider phone usage–related data as personal and sensitive. In selected developing economies, a high number of consumers, over 30 percent, use their phones to make financial transactions, and over 80 percent consider emails, calls, and texts as personal and sensitive data, while over 70 percent consider financial and medical data as private. Interestingly, in order to get a loan more easily, 70 percent of consumers would share data on mobile-phone usage and bank accounts, but only 60 percent would be willing to share data on social-media activity. Overall, as for the 2014 GSMA survey, 80 percent believe that the existence and adoption of good, clear policies and procedures governing data privacy would increase their trust in a financial institution.

Finally, the recommendations of the March 2017 ITU-T Focus Group on Digital Financial Services also noted consumer concerns about the sensitivity of their financial information.[18] The research referred to by the ITU indicates that consumers have concerns about how their information will be used and shared, and fear that it may expose them to "identity theft, embarrassment, and tax or criminal liability."[19] However, the report also notes that consumer attitudes vary from country to country.[20]

Although consumers' behaviors and attitudes toward the sharing of information vary from one jurisdiction to another, some trends can be highlighted. The above surveys suggest the following trends: (i) Consumers generally see data related to the usage of financial services and phones as sensitive. (ii) Nevertheless, they will not necessarily check privacy policies and forms of consent before accessing DFS. Finally, (iii) there is a demand for clarity and consistency in applicable rules and policies.

[15] *BCG Global Consumer Sentiment Survey, 2013* (Boston Consulting Group, 2013). The survey was conducted in five European countries (Germany, France, Italy, Spain, and the United Kingdom).

[16] "Mobile Privacy: Consumer Research Insights and Considerations for Policymakers" (GSMA, February 2014), available at **http://www.gsma.com/publicpolicy/mobile-privacy-consumer-research-insights-and-considerations-for-policymakers.**

[17] See, generally, the overall results of the survey in "Mobile Privacy."

[18] ITU-T Focus Group on Digital Financial Services, "Consumer Experience and Protection Recommendations," chapter 5 of *ITU Focus Group Digital Financial Services: Main Recommendations* (ITU, 2017), available at **http://www.itu.int/en/ITU-T/focusgroups/dfs/Documents/201703/ITU_FGDFS_Main-Recommendations.pdf**

[19] ITU-T Focus Group on Digital Financial Services.

[20] ITU-T Focus Group on Digital Financial Services.

4 HOW USAGE OF NEW FORMS OF DATA IN FINANCIAL SERVICES BRINGS BENEFITS TO CONSUMERS

This section analyzes the different ways in which new forms of data have been used and how they have been beneficial to consumers of financial services. New types of data, in particular big data, and the related analytics provide financial services firms with new opportunities to use greatly expanded data sets, and to combine historical data with real-time data when designing and marketing financial products. The potential benefits are numerous. Examples include the following: use of big data by financial services providers for client profiling;[21] putting in place better market-segmentation practices; assessing credit risk, including scoring models;[22] identifying and mitigating risks and delivering more tailored products;[23] assessing and preventing fraud in insurance claims;[24] and, finally, facilitating compliance with regulatory requirements, including meeting KYC compliance requirements.[25]

As mentioned, the increased uptake of mobile phones and electronic financial services is generating an enormous amount of data. Not only does this provide new multiple data sources, but the large volume enables big-data analytics that in turn can be used to identify consumer behaviors better and to monitor and prevent fraud. As identified by several reports and studies,[26] electronic-payment transactions have increased substantially over the years. This trend is expected only to increase as national authorities, governments, and regulators act proactively to increase the usage of electronic payments (for example, through the digitization of certain large-volume payment streams) and financial inclusion.[27] For example, banks in Indonesia and Hong Kong use predictive modeling to identify potential fraud and alert consumers almost instantly.[28] The analysis looks at either the "purchase-related data" or customers' habits and behaviors.

This phenomenon further allows markets and consumers to be analyzed ever more precisely, using millions of data points, and for increasingly tailored financial services and prices to be offered to consumers. For example, financial services providers can use all this new data to assess the product needs of a consumer, his or her risk profile and financial means, and even to establish whether a consumer may be willing to pay more for a given product than other consumers. Products and services can thus be designed to suit the needs of individual customers better, and offerings can be priced on an individual-risk basis.

Further, one of the most frequently cited and common advantages brought by analytics of this data is the possibility of developing credit scores for people who are not covered by traditional credit information systems, such as credit bureaus or credit registries. While access to credit "is a critical element of private sector–led growth,"[29] according to the Global Findex only 11 percent of the world's adult population borrowed formally in 2014 (this percentage drops to 9 percent in low-income countries, 8 percent in lower middle-income countries, and 10 percent in upper middle-income countries), and over 27 percent of firms globally identify access to finance as a major constraint.[30] This shows that many individuals and firms face significant constraint in accessing formal credit. This is partially due to the fact that many financial services providers are reluctant to provide credit due to poor infrastructure and limited coverage of credit information systems. Many individuals have no credit history. In fact, in countries where coverage is higher or where credit market infrastructure is stronger (on account of insolvency regimes, commercial courts, or movable collateral registries), more individuals and firms receive credit.[31] Given these con-

[21] Examples of new client-profiling technologies include Kopa Cash, which provides mobile phone–based loan approvals to M-Pesa account holders within minutes and advises on its website that social media is used in its operations. Nedbank makes widespread use of social-media analytics to enhance the experience of its banking customers.

[22] Examples include *Sesame Credit Management,* which creates credit scores for consumers and small businesses; *Kreditech,* which processes more than 20,000 data points per application using artificial intelligence built into private credit-scoring technology; *Cignifi,* which uses mobile-phone data, messages, and payments information to create credit scores; and *Lenddo,* which allows an organization to use its presence on social networks such as Facebook, LinkedIn, Google, Yahoo, and Twitter to prove its identity and creditworthiness.

[23] See **https://www.forbes.com/sites/bernardmarr/2015/12/16/how-big-data-is-changing-the-insurance-industry-forever/#7008b703289b**

[24] The South African insurer Santam has developed a new, faster system for processing claims based on big data to score claims for fraud risk.

[25] Tradle in the United States uses block-chain technology to bridge internal and external financial networks to accomplish user-controller KYC portability or Trulioo.

[26] See "Global Payment Systems Survey 2015: Accounts and Access" (World Bank Group, 2016), available at **http://pubdocs.worldbank.org/en/504871475847684346/GPSS-UFA-Note-October2016.pdf**

[27] See, generally, Committee on Payments and Market Infrastructures, "Payment Aspects of Financial Inclusion" (World Bank Group, September 2015), section 3.1.2.4, available at **http://www.bis.org/cpmi/publ/d133.pdf.**

[28] "Leveraging Data and Analytics for Customer-Centricity and Innovation" (The Asian Banker, April 2013), section 3, available at **https://www.theasianbanker.com/assets/media/dl/whitepaper/SAP_WP_2013_1.pdf**

[29] *Doing Business 2017: Equal Opportunity for All* (World Bank Group, 2017), 58.

[30] *Enterprise Surveys,* World Bank Group, available at **http://www.enterprisesurveys.org/.**

[31] See, generally, *Doing Business* 2017, 52–64.

straints and issues, the usage of big data for credit-scoring purposes represents an opportunity to provide credit to unserved and underserved. Many financial services providers and FinTechs—in developing economies and emerging markets[32] as well as high-income countries[33]—are turning to these new data sets and their analytics to develop credit and offer different types of credit products to financial services consumers to address information asymmetry–related issues.

Many financial services providers are also using big data to identify consumers and meet regulatory requirements. Around 18 percent of respondents to the Global Findex 2014 survey cited lack of necessary documentation as a barrier to having a formal account.[34] While there are many costs associated with offering financial services, including delivery costs in rural and remote areas, compliance costs also play a role in making delivery more expensive. In order to comply with regimes aimed at combating money laundering and the financing of terrorism, including KYC requirements, and to lower compliance costs, many financial services providers are using big-data tools.[35]

5 INTERNATIONAL STANDARDS RELEVANT TO BIG DATA AND FINANCIAL CONSUMER PROTECTION

This section outlines aspects of key FCP standards, including data-privacy and credit-reporting principles, that are applicable or relevant in a context where multiple new sources of data are being utilized and processed. The principal standards considered are (i) the G20 High-Level Principles on Financial Consumer Protection (G20 FCP HLPs);[36] (ii) the 2017 edition of the World Bank's *Good Practices for Financial Consumer Protection* (Good Practices);[37] (iii) the G20 DFI HLPs;[38] and (iv) the 2011 International Committee on Credit Reporting's *General Principles for Credit Reporting* (General Principles).[39] Reference is also made to data-protection standards where applicable, such as the 2013 OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (OECD Guidelines);[40] the 2009 International Conference of Data Protection and Privacy Commissioners' Madrid Resolution;[41] the Asia-Pacific Economic Cooperation (APEC) Privacy Framework;[42] and the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (Convention 108) of the Council of Europe.[42]

It is important to underline that there is not a single set of international standards, guidelines, and best practices on FCP. Rather, there are various international standards, guidelines, and codes that apply either generally or to specific issues or sectors.[43] Nevertheless, the standards referred to above cover the issues that seem most relevant in a context where multiple new sources of data, beyond traditional credit-reporting frameworks, are being used and processed.

Finally, it is important to underline that such principles may often be difficult to implement, particularly in low-ca-

[32] See for example, Catherine Cheney, "How Alternative Credit Scoring Is Transforming Lending in the Developing World," Devex, September 8, 2016, available at **https://www.devex.com/news/how-alternative-credit-scoring-is-transforming-lending-in-the-developing-world-88487.**

[33] For example, see Christina Farr, "Kabbage Brings Its Quick Fix Loans to UK Merchant," *VentureBeat*, February 16, 2013, available at **https://venturebeat.com/2013/02/16/kabbage-brings-its-quick-fix-loans-to-uk-merchants/.**

[34] Asli Demirguc-Kunt et al., "The Global Findex Database 2014: Bringing the 2 Billion Unbanked into the Formal Financial System," *Findex Notes*, no. 2014-2 (April 2015), available at **http://pubdocs.worldbank.org/en/113791483565360488/N2UnbankedV5.pdf**

[35] **http://financialservices.mazars.com/the-use-of-big-data-tools-to-improve-the-effectiveness-for-amlcft-and-kyc-policy/**.

[36] Available at **https://www.oecd.org/g20/topics/financial-sector-reform/48892010.pdf.**

[37] Available at **https://openknowledge.worldbank.org/handle/10986/28996?locale-attribute=en**

[38] Available at **https://www.gpfi.org/sites/default/files/G20%20High%20Level%20Principles%20for%20Digital%20Financial%20Inclusion.pdf**

[39] Available at **http://documents.worldbank.org/curated/en/662161468147557554/General-principles-for-credit-reporting**

[40] Available at **http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm**

[41] Available at **https://icdppc.org/wp-content/uploads/2015/02/The-Madrid-Resolution.pdf**

[42] Available at **https://rm.coe.int/1680078b37**.

[43] For example, see **International Standards on the Protection of Personal Data and Privacy: The Madrid Resolution** (International Conference of Data Protection and Privacy Commissioners, November 2009); **APEC Privacy Framework** (Asia-Pacific Economic Cooperation, 2005); ***General Principles for Credit Reporting*** (World Bank Group, September 2011), prepared by a task force coordinated by the World Bank; and **Responsible Digital Payments Guidelines** (Better Than Cash Alliance, July 2016). Other examples include Committee for Payments and Markets Infrastructure and the World Bank Group, **"Payment Aspects of Financial Inclusion"** (Bank for International Settlements, April 2016); ***Global Standard-Setting Bodies and Financial Inclusion: The Evolving Landscape*** (Global Partnership for Financial Inclusion, March 2016); and Basel Committee on Banking Supervision, **"Guidance on the Application of the Core Principles for Effective Banking Supervision to the Regulation and Supervision of Institutions Relevant to Financial Inclusion"** (Bank for International Settlements, March 31, 2016).

pacity environments and/or where changes in the financial sector are fast-paced. The principles have been developed by different organizations and multilateral bodies and provide a framework to ensure privacy and data protection. However, not only have many jurisdictions transplanted the frameworks into their national legal frameworks in different ways, but key aspects of the principles are difficult to implement effectively, particularly in environments where institutional capacity is low. As shown in section 6, these difficulties are heightened in a context where multiple new data sources are being used and processed.

5.1 Core Financial Consumer Protection Principles

Provided below is a brief high-level synthesis of FCP standards potentially relevant to big data.

- **Legal and regulatory framework.** Firstly, it is important that a proportionate, risk-based FCP legal and regulatory framework is in place that both applies to all service providers equally and covers key consumer protection issues. The G20 DFI HLPs refer to the need to "establish a comprehensive approach to consumer and data protection that focuses on issues of specific relevance to digital financial services" (Principle 5). The G20 FCP HLPs also emphasize the need for FCP to be an integral part of the legal framework (Principle 1). The 2017 edition of the Good Practices provides further details for the different types of DFS that are covered.

- **Institutional mandate and resources.** It is important that the relevant regulators have both a clear, adequate, and non-overlapping mandate (or at least that there is coordination if they are overlapping) and the required financial-sector skills and expertise, as well as the necessary capacity, tools, and resources. Further, the relevant FCP regulator should have formal mechanisms for consultation and coordination with other relevant regulators on policy and supervision issues (including privacy, telecommunications, and financial sector–specific regulators). Both the G20 FCP HLPs (Principle 1) and the 2017 edition of the Good Practices stress these points.

- **Disclosure and transparency.** Consumers should be provided with accurate, simple, and clearly expressed information about the features, risks, terms, and prices of financial services (including how their personal information will be handled). Both the G20 FCP HLPs (Principle 4) and the 2017 edition of the Good Practices confirm these points. These FCP standards further establish that such information should be given in a timely and user-friendly manner and that, where possible, the format should be standardized through the use of key summary documents. In a context where multiple types of data are being collected and processed, it becomes especially important that data-privacy and data-sharing policies are disclosed along with product features and comparability, as the more such data is used to provide personalized offers, the more difficult it will be for consumers to compare products.

- **Fair treatment and business conduct.** Consumers and investors should be treated fairly and respectfully; they should not be the subjects of discriminatory, misleading, or abusive treatment. Both the G20 FCP HLPs (Principle 3) and the 2017 edition of the Good Practices refer to these standards. The 2017 edition of the Good Practices also provides that unfair terms and practices should be prohibited and closely monitored by the relevant authority. In a context where several different types of new data are being collected and processed, these standards are likely to mean that there should be a focus on the terms of privacy consents and on company policies and procedures followed when dealing with personal information. When large volumes of data are collected and used for a wide variety of purposes, it is especially important that such standards are in place to protect consumers with limited financial capability and technological capacity.

- **Product suitability.** Products should be designed with consumer needs in the target market in mind, and the financial needs and capacity of a consumer should be considered before a financial service is provided. The G20 FCP HLPs (Principle 6) and the 2017 edition of the Good Practices reflect these standards. While new multiple forms of data allow for more tailored and, hence, more suitable products to be provided, it is important that the data used is relevant and accurate and that the product meets the needs and capacity of the consumer. This is particularly important in the context of credit products: A credit facility should be provided only after a reasonable assessment is made of the consumer's ability to repay it. Further, this is even more important when new types of data, beyond traditional credit reporting, are used not to provide products that are suitable to users, but to maximize profits.

- **Customer mobility.** Consumers should be able to transfer a financial facility to a new provider without unreasonable difficulty. In light of a growing international focus on the issue, the 2017 edition of the Good Practices specifically provides that regulators should ensure that rules forbid anti-competitive practices and enable customer mobility. To facilitate such mobility, it is especially important that consumers are able to request the transfer of their personal information from one financial services provider to another, and that the information is transferred in a useful format.[44]

- **Privacy and data protection.** A consumer's personal information should be kept confidential and secure, and used only if accurate and with consent (subject to applicable law). The G20 FCP HLPs (Principle 8) and the 2017 edition of the Good Practices cover this aspect of data protection and privacy. The G20 FCP HLPs also refer to the need for consumers to be informed about data-sharing, access, and correction rights and to have inaccurate or unlawfully collected data deleted. The revised Good Practices also provides guidance on the lawful collection of data, the usage and storage of such data, and sharing such data with third parties. A more detailed analysis of key data-privacy principles is provided below.

5.2 Privacy and Data Protection

As mentioned above, beyond the core FCP principles, more specific and important data-protection principles are directly relevant to financial consumers in a context in which multiple data sources are being collected and used. Below is a summary of the key data-privacy principles:

- **Informed and meaningful consent.** Consent is a fundamental principle concerning data privacy and FCP. Different standards, however, define and deal with consent differently. The OECD Data Collection Principle provides that "there should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject."[45] Similarly, the OECD Use Limitation Principle refers to the need for consent if data is to be used for purposes other than the original purpose of collection.[46] The Madrid Resolution, through its General Principle of Literacy, provides that, as a general rule, data processing should occur only "after obtaining the free, unambiguous and informed consent of the data subject" (subject to limited exceptions).[47] The APEC Privacy Framework also refers to the need for consent to the collection of personal information (where appropriate).[48] Finally, while Convention 108 does not include a similar principle, it goes beyond the consent principle, specifying that data undergoing automatic processing shall be stored safely and for a legitimate purpose, forbidding any other use beyond that specific purpose.[49] Convention 108 further specifies that certain categories of sensitive data cannot be processed automatically, regardless of whether the user has given consent or not, unless national legislation provides appropriate safeguards. The relevant categories include "personal data revealing racial origin, political opinions or religious or other beliefs, as well as personal data concerning health or sexual life."[50]

- **Security of data.** Another key principle relating to data privacy and FCP is the need for financial consumers' data to be kept safe and secure. The OECD Security Safeguards Principle provides for security safeguards as follows: "Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data."[51] The Madrid Resolution places stronger emphasis on security issues, containing separate principles on security measures and the duty of confidentiality.[52] In particular, there is an obligation to protect personal data with the appropriate technical and organizational measures to ensure the data's integrity, confidentiality, and availability. There is also an obligation to inform data subjects of significant data breaches. Similarly, Convention 108 also provides for "appropriate security measures against accidental or unauthorised destruction or accidental loss as well as against unauthorised access, alteration or dissemination."[53] The APEC Privacy Framework contains similar requirements (though with less details) in Principle VII, "Security Safeguards."

- **Access, rectification, correction, and opposition (ARCO) rights.** Another key issue relating to data privacy, FCP, and new types of data is consumers' right to access the data collected and to request rectification and cancellation of such information. To this extent,

[44] The EU General Data Protection Regulation to this extent also introduces the concept of data portability, see Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC, Recital 20: "(1) The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided, where: a) the processing is based on consent pursuant to point (a) of Article 6(1) or point (a) of Article 9(2) or on a contract pursuant to point (b) of Article 6(1); and b) the processing is carried out by automated means."

[45] Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data (OECD, 2013), Principle 7.

[46] OECD Guidelines, Principle 10.

[47] Madrid Resolution, Principle 12(a).

[48] APEC Privacy Framework, Principle 18.

[49] Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (Council of Europe, 1981), Article 5(b).

[50] Convention 108, Article 6.

[51] OECD Guidelines, Principle 11.

[52] Madrid Resolution, Principles 20 and 21.

[53] Convention 108, Article 7.

the OECD Individual Participation Principle, in summary, provides consumers with the right to find out if a data controller[54] has information about them; to obtain that data within a reasonable time, in an intelligible form, and at a charge that is not excessive; and to challenge the data and, if successful, to have the data erased or corrected.[55] The Madrid Resolution contains similar principles dealing with the right of access and rights to rectify and delete.[56] Further, the Madrid Resolution Data Quality Principle provides that when data is no longer necessary for the legitimate purposes of collection, then it should be deleted or rendered anonymous.[57] The additional safeguards of Convention 108 provide for similar rights and also specify that data used for a purpose not allowed by law needs to be erased.[58] The APEC Privacy Framework has detailed provisions on access and correction rights in Part V. There are, however, broad exceptions to these general principles. They include the right to refuse a request where granting access would involve unreasonable business expense, where the information should not be disclosed because of legal or security reasons, or where access could jeopardize confidential commercial information or violate another individual's privacy.[59]

ARCO rights have also been covered by international principles that go beyond data protection, including key financial sector–related principles, as they are essential data-protection rights. As highlighted in the recent G20 DFI HPLs, "Consumers also need to have transparent, affordable and convenient access and correction rights."[60] ARCO rights are especially relevant in a DFS context when an individual's data is held, or can be accessed, by multiple institutions and the data may be in many different forms.[61] Consumers may not know who is holding, or has access to, their data,[62] for what purpose it is being used, where it is being held or by whom, or the nature and scope of the data that is being held. And even if individuals do know all this information, they are not likely to be able to enforce those rights, especially where customer-recourse systems are not clearly stated, and especially where data is held in the cloud and/or is unstructured data. It is also important that data can be accessed in a usable form.[63]

- **Accuracy and reliability of data.** The OECD Data Quality Principle (Principle 8) provides that data collected should be relevant to the purpose for which it is to be used and should be "accurate, complete and kept up-to-date." The APEC Privacy Framework has similar provisions (Principle 21). The Madrid Resolution Data Quality Principle also requires that data always be accurate and sufficient and kept up to date. Like other principles, Convention 108 requires that data files be accurate and up to date (Article 5).

5.3 Credit Reporting

Finally, to analyze principles relevant to FCP and big data completely, it is important to look at credit reporting and the applicable principles. The General Principles address the use of credit information and other relevant data in credit-reporting systems. These systems involve the collection of credit data and other relevant information from multiple sources, resulting in the elaboration of credit reports or other products that are based on data analytics, such as credit scores. Figure 1 summarizes the principles and gives guidance for interpreting each principle when big data is being used for making risk-based credit decisions. The General Principles contain provisions allowing providers of credit-reporting services to collect and process "all the relevant information needed to fulfill their lawful purposes," and requiring data to be kept, to the extent possible, "free of error, truthful, complete and up to date."[64]

Applying these principles when using new types of data beyond credit reporting to evaluate individuals' creditworthiness may be challenging. Particularly when developing credit scores, recent trends show that new credit risk/credit-scoring tools make extensive use of big data.[65] This new trend brings new opportunities to consumers who have no credit history or "thin files," making them unserved or underserved. Figure 2 shows the develop-

[54] *Data controller* has been defined as "a party who, according to national law, is competent to decide about the contents and use of personal data regardless of whether or not such data are collected, stored, processed or disseminated by that party or by an agent on its behalf." The OECD Privacy Framework (OECD, 2013), chapter 1, annex.

[55] OECD Guidelines , Principle 11.

[56] Madrid Resolution, Principles 16 and 17.

[57] Madrid Resolution, Principle 9.

[58] Convention 108, Article 8.

[59] APEC Privacy Framework, Principles 23–25.

[60] See the G20 DFI HLPs, Action Item, Principle 5.

[61] Again, this raises the issue of who should be considered the "data controller" in this context.

[62] Again, this raises the issue of who should be considered the "data controller" in this context.

[63] Omer Tene and Jules Polonetsky, "Big Data for All: Privacy and User Control in the Age of Analytics," *Northwestern Journal of Technology and Intellectual Property* 11, no. 5 (2013), 239.

[64] *General Principles,* Principle 1.

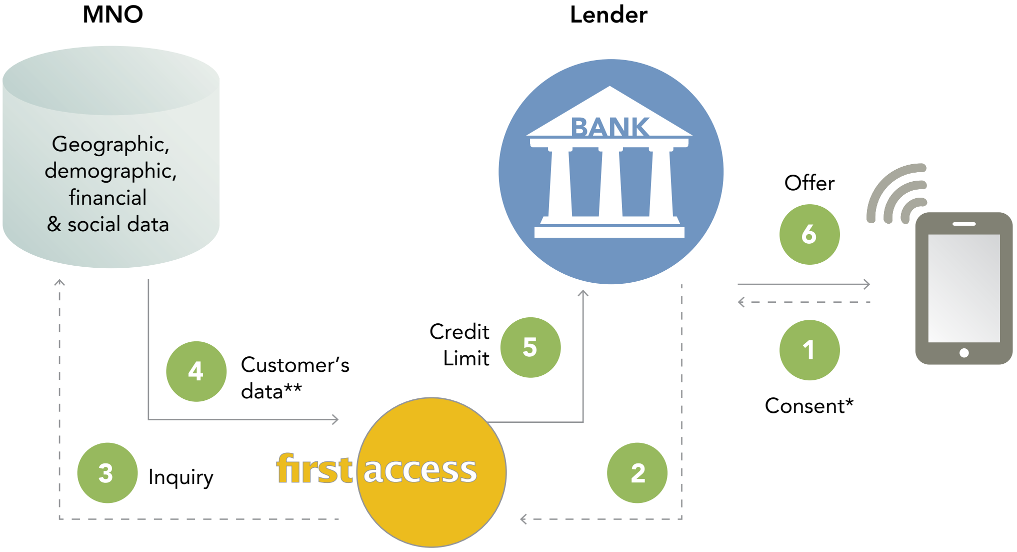
[65] For the purpose of credit scores to evaluate creditworthiness of individuals, big data includes information generated by traditional business activities and from new sources—such as electronic-payments data from point-of-sale terminals, bank automated-teller machines, mobile-network operators, and utilities—combined with social media (Facebook and Twitter posts and YouTube videos) and geodemographic data.

FIGURE 1: The General Principles for Credit Reporting

GPI Data	GPII Data processing: security and efficiency	GPIII Governance and risk mangement	GPIV Legal and regulatory environment	GPV Cross-border data flows
Credit reporting systems should have accurate, timely and sufficient data—including positive—collected on a systematic basis from all reliable appropriate and available sources, and should retain this information for a sufficient amount of time	Credit reporting systems should have rigorous standards of security and reliability, and be efficient	The governance arrangements of credit reporting service providers and data providers should ensure accountability, transparency and effectiveness in managing the risks associated with the business and fair access to the information by users	The overall legal and regulatory framework for credit reporting should be clear, predictable, non-discriminatory, proportionate and supportive of data subject and consumer rights. The legal and regulatory framework should include effective judicial or extra-judicial dispute resolution mechanisms	Cross-border credit data transfers should be facilitated, where appropriate, provided that adequate requirements are in place

Source: Own elaboration based on the General Principles.

FIGURE 2: Mobile Scoring Methodology



* Customer consents to sharing mobile usage data
** Usage for given period shared one

Source: First Access, World Bank Financial Infrastructure Week, 2015.

ment of scoring models using nontraditional data captured from different sources, and a discussion of the relevant principles is provided below.

- **Data.** Data used for credit-reporting principles needs to be reasonably accurate. Considering that in this context, data could be pulled from a wide variety of sources, General Principle 1 is extremely relevant. To cover potential risks of inaccuracy, it provides that "credit reporting systems should have relevant, accurate, timely and sufficient data—including positive—collected on a systematic basis from all reliable, appropriate and available sources, and should retain this information for a sufficient amount of time."[66] This principle could be applied not only to credit-reporting systems per se but also to providers of credit-information services, such as Fintech companies developing scores, or lending platforms with embedded scoring models in their systems. Accuracy problems emerge when small and big amounts of structured and unstructured data are pulled from multiple sources and the information is not updated on a systematic basis from the same sources. Also, some of this information will be self-reported by the data subjects. When all this information is merged and used to develop credit scores, the following questions arise:
  - Lawfulness of data collection and compatible purposes, and the role of consent when the purpose of data collection is neither specifically covered under the law nor compatible with the original purpose of data collection.
  - How to define and ensure enforcement of data-retention periods.
  - In the case of open sources such as application programming interfaces (API), would this information be considered public information?
- **Data-processing security and efficiency.** Like data privacy–related principles, the General Principles also cover data security but expand on the security of the processing of data and its efficiency. The relevant principle establishes that "credit reporting systems should have rigorous standards of security and reliability and be efficient and protect data against any loss, corruption, destruction, misuse or undue access" (General Principle 2). Although difficult, it might be necessary to assign a line of accountability when data is pulled from many different sources and captured through a dynamic database. In this case, data-sharing agreements between the different parties might be necessary. Evidence of consumers' choices regarding the collection, processing, and further use of their data might be needed as well. Also, the number of vulnerabilities increases as the number of different players grows, all of whom have different technologies, networks, software, and security policies for access to and use of such data. Finally, as the data chain increases and number of players becomes larger, it becomes difficult to address vulnerabilities and to develop communications and an adequate remedy protocol. Implementing a right to object to the use of information for certain purposes might require specific guidelines enabling the adoption by at least key data sources.
- **Governance and risk management.** In this analysis, another important aspect of credit reporting is the relevant governance arrangements. In fact, when evaluating creditworthiness of individuals, "governance arrangements for credit reporting service providers and credit reporting data providers should ensure timely and accurate disclosure of relevant matters related to the entity and its activities" (General Principle 3). This issue includes (i) the legal framework supporting the activities of such institutions, (ii) the types of entities that may become users of the scores developed and the conditions under which they use such services, (iii) the rules and procedures for collecting and processing such data, (iv) the uses of data, and (v) mechanisms for identifying and mitigating risks involved in the activity. In addition, another transparency rule relates to the protocol established to dispute errors in data used to develop the scores. In traditional credit-reporting systems, it is easy to identify and assign accountability to the data controllers. In an environment where data is pulled from different sources in a dynamic manner, this chain of accountability becomes more opaque.
- **Legal and regulatory environment.** The big-data effect has often raised concerns about the potential risk to consumer privacy when data is combined, resulting in precisely tailored consumer profiles. The data sets may be vast, but they can be used to identify individual needs, habits, and financial patterns accurately. Consumers, however, are often unaware that they are generating data that affects analytical models. General Principle 4, on the legal and regulatory framework, includes guidelines on consumer rights. Rules regarding the protection of data subjects/consumers should be clearly defined. At a minimum, these rules should include ARCO rights defined as "(i) the right to object to their information being collected for certain purposes and/or used for certain purposes, (ii) the right to be informed on the conditions of collection, processing and distribution of data held about them, (iii) the right to access data held about them periodically at little or no cost, and (iv) the right to challenge accuracy

[66] *General Principles,* Principle 1.

of information about them."[67] The General Principles recognize consumers' consent when third parties access their data although the GPs also call for the collection of information—including positive data—in a comprehensive manner. In an open environment where the data controller is difficult to identify and the purpose of data use might differ completely from the purpose of data collection, consumers' choice regarding his/her own data might become more necessary and consent mechanisms, coupled with the concept of portability could provide certain protection to consumers when third parties use extensively their information.

- **Crossborder data flows.** Finally, considering the DFS context, crossborder data flows are relevant. The General Principles establish that "cross-border flows should be facilitated provided that specific requirements are in place" (General Principle 5). In fact, while it is important to allow for crossborder data flows, it is important that adequate frameworks are in place to ensure that consumers' rights are adequately protected. Crossborder flows are even more relevant in a context where multiple new sources of data are being utilized; it is important that adequate requirements are in place to protect the processing of such data. This principle is aimed at facilitating the flow of credit information across borders when market conditions call for such flow. However, new methodologies developed to evaluate creditworthiness of individuals and small and medium-sized enterprises already involve the processing of data in different jurisdictions where the service is provided or for the consumers' data to be captured and further exploited.
  - The number of tools used to collect data are far more numerous.
  - The number of jurisdictions involved in the data processing is also larger.
  - Discussion of cloud computing and the need to identify a specific jurisdiction is intimately linked to the identification of a data controller.

6 FINANCIAL CONSUMER PROTECTION PRINCIPLES, NEW FORMS OF DATA CHALLENGES, ISSUES, AND RISKS

While the benefits of using and processing new types of data can be substantial, the risks for consumers of financial services can be consequential. New data, beyond traditional credit-reporting frameworks, allows for the delivery of more tailored financial services to consumers and has great potential to increase access to financial services, but its use can also pose FCP challenges, issues, and risks and can introduce new challenges to the application of the standards discussed. This section provides an analysis of these issues, with practical examples provided where relevant.

6.1 Consent

The question at the center of concerns around consent in a big-data context is, How can informed consent be given effectively? Consent is especially problematic when the intended purpose for using the data is unknown at the time of collection or is constantly changing and expanding. Another barrier to giving effective consent is the likelihood in a financial-inclusion context of high levels of illiteracy and low levels of financial capability, especially where deemed consent is provided for in lengthy terms and conditions. Another issue is whether local laws impose restrictions on giving consent to especially sensitive data, such as data relating to health, religious or political affiliations, or sexuality.[68]

Consent is especially important in the context where different types of data can be collected, used, and shared by different providers and across borders. An analysis of the terms and conditions of financial products conducted for this note shows that it is common practice to include clauses that allow for (i) the transfer of data to subsidiaries and partners (such as telecommunications companies that may be sharing the data with other financial services provider partners); (ii) the storage of personal information, including transactional, messaging, and call data, for up to five years but limiting providers' liability for data breaches; and (iii) the transfer of data from certain jurisdictions (not all jurisdictions, including ones with strong protection) to others for the purpose of processing such data.

The limitations on the consent model include:

- **Standard "adhesion" contracts.** Financial services contracts, and especially those formed digitally without any human interaction, are likely to be in a standard form. Consumers do not have the right to opt out of such contracts or to negotiate them in any way. Accordingly, the terms relating to personal information are usually provided on a "take it or leave it" basis.
- **Length and complexity.** Forms of consent are so long and complex that it is unrealistic to expect consumers to read or understand them.

[67] *General Principles,* Principle 4.

[68] See, for example, Ira S. Rubinstein, "Big Data: The End of Privacy or a New Beginning?," *International Data Privacy Law* 3, no. 2 (2013).

- **Incomplete information.** All the information needed to make an informed decision may not be provided to consumers.[69]

- **Hidden forms of consent.** The terms of a privacy consent are often unclear and hidden within lengthy terms and conditions, implying that the user is unaware that consent to the collection has even been given.

- **Multiple consents.** Consumers may be expected to consider more than one form of consent when acquiring a product. For example, a corporate privacy policy, a consent for the product provider, and a consent for any related partner (such as an e-wallet provider or a partner bank providing credit) may all need to be considered.

- **Lack of choice.** Consumers do not have a meaningful choice if they want the product in question. Behavioral research shows that while consumers are interested in the way their data could be used for loan determination, their overall need for a loan would supersede concerns for privacy.[70]

- **Opt out, rather than opt in.** Use of information for, say, marketing purposes may be permitted subject to an opt-out, rather than an opt-in, right, and consumers may not even be aware that they have this right.

- **Consent after the event.** Consent may be requested after the information has already been used. For example, social-media and air-time information may be used to create a credit score that is used as the basis for an unsolicited offer of credit.

- **No expiry date.** There is usually no expiry date to the consent, and consumers cannot be expected to maintain the same preferences over time (especially given rapid technological developments in big data) or even to remember what they have agreed to.[71]

Limitations on the consent model are especially acute in a financial-inclusion context. In some countries, consumers are increasingly being asked to provide wide-ranging consent to the use of personal information derived from big-data sources with limited understanding. ITU research shows that in African countries, vague, complex, and buried consent clauses are fairly common. In fact, 83 percent of the "contracts reviewed had clauses that permit the provider to share information with third parties, such as credit reference bureaus, law enforcement agencies (both domestic and international), regulators, provider agents, lawyers, auditors, and subsidiaries . . . [including] for reasonable commercial purposes related to the provision of services."[72] However, their ability to do so effectively is more likely to be challenged by illiteracy, an inability to understand complex and lengthy terms relating to the collection of multiple types of data, and the printing of the consent terms in a language the consumer does not understand (for example, in English in a country that has multiple local languages). Further, such consumers are more likely to be pressured into feeling that they have no choice but to agree to any form of consent if they want the product.

6.2 Product and Price Segmentation and Potential Discrimination

While the collection, usage, and processing of new types of data allow for better tailoring of financial services and products, if the above-mentioned principles are not adequately respected, use of the data may lead to unfair product and price segmentation (including discrimination). As mentioned above, the FCP principles call exp ressly for a prohibition of unfair and abusive practices, including discriminatory ones. Similarly, data-privacy standards call for limitations on data collection, including in some instance forbidding the usage of specific types of sensitive data, and requirements to tie data to the specific purpose for which it was collected.

Various concerns have been expressed about the use of multiple new types of data to analyze and segment consumers, including:

- **Discrimination.** There is the potential to be able to use big data to discriminate against consumers with particular attributes, whether inadvertently or deliberately. The U.S. Federal Trade Commission put the issue another way, asking, "Big data: a tool for inclusion or exclusion?"[73] One of the most fundamental human rights, and FCP principles, is that a person should not be unfairly discriminated against. The key point is that discrimination on any ground—including race, color,

[69] It is generally held that the information that should be provided includes which data are collected, used, and shared; the purposes for which data are used; which security measures are taken; who is processing the data and who is accountable; and user rights and how they can be exercised. Bart Custers, "Click Here to Consent Forever: Expiry Dates for Informed Consent," *Big Data & Society,* January–June 2016: 1–6, available at **http://journals.sagepub.com/doi/10.1177/2053951715624935**.

[70] See, generally, Costa, Deb, and Kubzansky, 2016.

[71] Custers, 2016.

[72] ITU-T Focus Group on Digital Financial Services, "Review of DFS User Agreements in Africa: A Consumer Protection Perspective" (ITU, January 2017), available at **https://www.itu.int/en/ITU-T/focusgroups/dfs/Documents/01_2017/ITU_FGDFS_Report-on-Review-of-DFS-User-Agreements-in-Africa.pdf**

[73] See, generally, "Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues" (Federal Trade Commission, 2016), available at **https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf**

BOX 1

Emerging Approaches to Address Consent-Related Issues

Given the limitations in the consent model, alternatives to the need for effective and informed consent, and innovative ways to obtain consent, are being widely discussed and, to some extent, implemented. Below are key examples.

- **Privacy by design.** Put simply, this concept envisages building privacy into all stages of the design and architecture of information systems, business processes, and networked infrastructure. The focus is on taking a proactive, preventive approach to the protection of privacy and the avoidance of privacy harms. Created by the then information and privacy commissioner of Ontario, the concept rests on the following seven principles:
  1. Proactive, not reactive; preventive, not remedial
  2. Privacy as the default setting
  3. Privacy embedded into design
  4. Full functionality—positive-sum, not zero-sum
  5. End-to-end security—full life-cycle protection
  6. Visibility and transparency—keep it open
  7. Respect for user privacy—keep it user-centric[74]
- **The EU General Data Protection Regulation.** This is the leading example of a regulatory approach that requires privacy-by-design principles. According to Article 25, "The controller shall . . . implement appropriate technical and organisational measures . . . in an effective way . . . in order to meet the requirements of this Regulation and protect the rights of data subjects."
- **Minimal collection of information.** This is the concept that only the minimal amount of data should be collected. The General Data Protection Regulation has also implemented this principle. Article 5(1)(c) provides: "Personal data shall be . . . adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')."
- **Tiered consents.** It may be more appropriate to introduce a concept of tiered consent by which consumers will be required to give different types of consent for the processing of certain types of data or for specific purposes.
- **An expiry date for consents.** Suggestions have been made that, given that consents are virtually never reviewed or renewed, there might be a limitation period on the effectiveness of some forms of consent. It is, however, acknowledged that such an approach will not solve all the issues with informed consents.[75]
- **An opt-in, rather than an opt-out, consent.** The recitals to the General Data Protection Regulation state that "silence, pre-ticked boxes or inactivity should not therefore constitute consent."[76]
- **Using simple messages to help consumers understand which data is being collected and for what purposes.** Research by the Consultative Group to Assist the Poor has highlighted that simple messages delivered via SMS may help customers understand concepts relating to the collection of new types of data and encourage them to consider data-protection issues. However, the research also showed that consumers are prepared to allow use of their data if it means that they can obtain a loan.[77]

Digital technology could also be used to facilitate the taking of informed consent. As seen in section 3, while many consumers care about giving meaningful consent, they often provide it without reading the terms and conditions of their consent. To address these issues, consideration could be given to developing tools that provide for simpler, more clearly expressed, and highlighted forms of consent. Such tools could well be technology-based. They could include a requirement for the use of standardized forms of consent, as well as the option of having verbal forms of consent that would be recorded by the financial services provider.

[74] Ann Cavoukian, "Privacy by Design: The Seven Foundational Principles" (Information and Privacy Commissioner of Ontario, 2011), available at https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf.

[75] Custers, 2016.

[76] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC, Recital 32.

[77] Rafe Mazer, Jessica Carta, and Michelle Kaffenberger, "Informed Consent: How Do We Make It Work for Mobile Credit Scoring" (Consultative Group to Assist the Poor, 2014), available at http://www.cgap.org/research/publication/informed-consent-how-do-we-make-it-work-mobile-credit-scoring

Source: Own elaboration based on different cited sources.

sex, language, religion, political or other opinion, national or social origin, property, birth, or other status—should not take place.[78] These principles have been reflected in equal opportunity laws globally, as well as in international standards concerning FCP. For example, in an action item for Principle 5, the G20 DFI HPLs state: "Require that data not be used in an unfair discriminatory manner in relation to digital financial services (e.g., to discriminate against women in relation to access to credit or insurance)."

- **Opaque algorithms that perpetuate biases and assumptions.** As discussed above, both key FCP principles and the General Principles provide that consumers should be aware of the policies and procedures used to assess creditworthiness and the needs of financial consumers. Commentators also argue that the algorithms used are opaque and may perpetuate biases and contain built-in assumptions that are not valid and may be unfair.[79]
- **Differential pricing or "price discrimination."** This is the practice of charging consumers' different prices for the same product, without reference to cost considerations or risk. The practice takes advantage of the willingness of some customers to pay more without losing other more price-sensitive customers.[80] The use of new types of data from multiple sources makes the practice much easier and cheaper, given big data's ability to provide increasingly segmented customer information. The concern is that it may result in unfair treatment of consumers (including financial consumers).
  - *Evidence from the United Kingdom.* A study conducted by the Financial Conduct Authority[81] shows that both price discrimination and cross-subsidies are common practices across the retail financial services market in the United Kingdom. The Financial Conduct Authority noted that there are potential competition concerns and that vulnerable consumers may be harmed by the practice while recognizing that price differentiation could also be beneficial. Further, while regulatory interventions on pricing need to be approached cautiously, the authority noted that "regulatory interventions that stop short of direct setting or capping of prices appear to have had positive impacts in addressing problems arising from price discrimination."[82] Examples cited of such interventions include the banning of joint credit and payment-protection insurance.
  - *The United States case:* Research conducted in 2015 by the White House and the Federal Trade Commission has indicated that the use of big data may result in discriminatory pricing.[83] This is because consumers tend to be associated with their network of friends, relatives, and ethnicity. As a result, only certain communities (in particular, African American communities) may be offered products at a higher price. Research conducted in 2016 by the Federal Trade Commission also highlighted the concern that "big data analytics could affect low-income, underserved populations, and protected groups" (especially in relation to credit and employment opportunities).[84] Other commentators have noted that, although there are arguments that algorithms can eliminate human biases, "an algorithm is only as good as the data it works with.".[85] The selection of key attributes used in algorithms is also relevant as search engines' algorithms may learn to prioritize characteristics associated with a group on individuals (e.g. minorities, women) more frequently than other characteristics not necessarily associated with those groups. Therefore, it might be useful to understand how meaningful are the correlations found by the analytics tools based on big data.

6.3 Comparability of Financial Services and Products

A fundamental FCP principle is having standardized disclosure for selected products and services to enhance comparability. However, the delivery of personalized services can limit, if not render impossible, comparability among providers with different levels of information about the consumer. Although the issue has not been discussed

[78] International Covenant on Civil and Political Rights (Office of the High Commissioner on Human Rights, 1976), available at **http://www.ohchr.org/EN/ProfessionalInterest/Pages/CCPR.aspx**.

[79] See, for example, Megan Smith, DJ Patil, and Cecilia Muñoz, "Big Risks, Big Opportunities: The Intersection of Big Data and Civil Rights," *What's Happening,* May 4, 2016, available at **https://obamawhitehouse.archives.gov/blog/2016/05/04/big-risks-big-opportunities-intersection-big-data-and-civil-rights**

[80] "Big Data and Differential Pricing" (Executive Office of the President, 2015), available at **https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/docs/Big_Data_Report_Nonembargo_v2.pdf**

[81] Pete Lukacs, Leslie Neubecker, and Philip Rowan, "Price Discrimination and Cross-subsidy in Financial Services," Occasional Paper no. 22 (Financial Conduct Authority, September 2016).

[82] Lukacs, Neubecker, and Rowan, 2016, 9.

[83] "Big Data and Differential Pricing" (Executive Office of the President, February 2015), available at **https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/docs/Big_Data_Report_Nonembargo_v2.pdf**, and "Big Data: A Tool for Inclusion or Exclusion?"

[84] "Big Data: A Tool for Inclusion or Exclusion?," executive summary.

[85] Solon Barocas and Andrew D. Selbst, "Big Data's Disparate Impact," *California Law Review* 104 (2016).

BOX 2

Insurance, an Area Where Big Data Could Become a Potential Source of Discrimination

Exceptions to discrimination laws generally allow insurers to discriminate based on actuarial or statistical data on which it is reasonable to rely. The usage of new types of data coming from multiple sources means that insurers can do this more efficiently. Being able to price insurance on a policyholder basis means low-risk customers do not have to subsidize high-risk customers, and adverse selection risks will be minimized. These risks can make insurance markets fail or lead to insurance not being provided because of the risk that insurers will be left with high-risk customers when low-risk customers self-insure. Insurers may also be able to price more effectively for moral hazard risk—the risk that insureds will take less care of the insured property because they have insurance to cover the risk. This practice will also better enable insurers to understand, and price, the actions that consumers can take to avoid risk—such as medical vaccinations for life insurance or watering crops for crop insurance policies. And they can help insurers monitor whether the required actions have in fact been undertaken.[86]

[86] See, generally, Max N. Helveston, "Consumer Protection in the Age of Big Data," *Washington University Law Review* 93, no. 4 (2016), available at **https://openscholarship.wustl.edu/law_lawreview/vol93/iss4/5/**

Source: Own elaboration based on cited sources.

widely,[87] the concern here is that the increased personalization of offers could limit consumers' ability to compare financial products and services.[88] In fact, while more tailored financial services can be beneficial to consumers, if this is based on data that only one or a certain group of financial services providers can access (for example, telecommunication/air-time data that only the e-money issuer subsidiary can access), consumers may not receive or have access to comparable information for the relevant product, as this is based on a particular set of data to which only certain providers can have access to.

Limited comparability and the possibility that certain providers may retain a monopoly on the information of certain consumers, or even categories of them, can ultimately hinder competition. While the overall goal of providing consumers with comparable information is to lower prices by increasing competition, if the offer of a relevant financial service is based on data to which only the provider has access, it can in turn have an adverse effect on competition and ultimately the prices that consumers pay. When this phenomenon relates to a large group, or even the vast majority of consumers, this can result in a monopoly, as only one provider will effectively be able to offer the service.

6.4 Security

Data security is clearly of prime importance in relation to personal data. However, it is not clear what minimum-security protocols would be appropriate (or how they could be enforced) in relation to this phenomenon where new types of data are being collected from multiple sources, as it is often unstructured and may be in multiple hands and jurisdictions. Innovations such as the India Data Locker, which allows for the secure storage of e-documents in personal lockers in the cloud, may be part of the solution but do not cover the entirety of an individual's personal information. Addressing security aspects in open environments, with a large number of institutions and users accessing such data, increases the level of difficulty. Further, the possibility of storing large amounts of data also increases potential; in fact, the potential fall out of a cyber-attack could have a long term impact if personal data is used is not used in the immediate term but in the future.

There are numerous examples of security weaknesses that affect consumers' data. Among them is the Equifax data breach that occurred between March and July 2017 in the United States and affected 143 million consumers included in the credit bureau. According to the Federal Trade Commission, hackers accessed "people's names, Social Security numbers, birth dates, addresses and, in some instances, driver's license numbers. They also stole credit card numbers for about 209,000 people and dispute documents with personal identifying information for about 182,000 people."[89] The General Principles include guidelines to avoid or at least mitigate the loss, corruption, destruction, misuse, or undue access of data. In this

[87] The three European financial-sector regulators have begun looking at this issue. See, generally, "Joint Committee Discussion Paper."

[88] See, generally, "Joint Committee Discussion Paper."

[89] Seena Gressin, "The Equifax Data Breach: What to Do," *Consumer Information*, September 8, 2017, available at **https://www.consumer.ftc.gov/blog/2017/09/equifax-data-breach-what-do**.

context, it suggests the adoption of measures and the frequent review of the adequacy of such measures. The objective of these safeguards should be to contain, limit, and respond to data-security breaches.

A key question that arises is whether traditional data-security processes and procedures are adequate. It is questionable whether traditional security methods—access codes, authorization levels, firewalls, and so forth—will still apply in a world of cloud computing and the collection of large amounts of structured and unstructured data. However, the vast scale of the different sources of data as well as the large volume of big data make it difficult to know what the minimum protocols to protect a consumer's personal data should be, especially his or her financial data.

While it is unclear whether the examined principles are adequate to address such issues, new principles and good practices are emerging. In particular, see the recent G20 DFI HLPs, where an example of an implementing action item is to "develop guidance to ensure the accuracy and security of all data related to: accounts and transactions; digital financial services marketing; and the development of credit scores for financially excluded and underserved consumers. This guidance should cover both traditional and innovative forms of data (such as data on utility payments, mobile airtime purchases, use of digital wallet or e-money accounts, social media and e-commerce transactions)."

6.5 Accuracy and Reliability of Data

The data-protection standards referred to above, as well as the General Principles, emphasize the importance of having accurate and reliable data. Implementing such principles when data is obtained from a wide variety of sources can pose significant challenges, as data accuracy and reliability may be harder to check. As seen in section 4, new types of data coming from multiple sources are generally used to assess creditworthiness when financial information about the borrower is absent either because he or she has recently entered the formal financial sector (for example, young people or newly arrived immigrants) or because the credit information system is not developed. The inherent risk with using inaccurate and/or unreliable data is that the score/worthiness assessment may be erroneous, leading to potential risks of exclusion if the error is in underestimating the ability of the borrower to repay, or to potential risks of overindebtedness if the error is in overestimating the ability of the borrower to repay.

New types of structured and unstructured data coming from multiple sources make accuracy-related issues more relevant. For example, opinions, intentions, and historical data may be collected, which can lead to problems in determining accuracy and currency, even though the data might be considered useful for marketing, product-development, and credit-assessment purposes. Different items of data may be combined for a particular purpose without any item in itself being considered to be "complete" for that particular purpose.

6.6 Crossborder Data Flows

Convenient online services that allow consumers to access products and services anywhere might also entail crossborder data flows. Building trust in the online environment is key and involves the collaboration of all participants, including consumers, data providers, service providers, and authorities. In addition, common rules for international cooperation are also relevant in order to achieve greater cooperation between authorities. Further consideration could be given to this aspect, as it involves the need for a harmonized international approach to consumers' rights, dispute-resolution mechanisms, accountability for data errors, and data-security measures. In addition to the challenges and risks cited under this section when implementing the General Principles and other data-privacy and FCP principles, the implementation of General Principle 5, which covers crossborder data flows, might also involve a conflicting set of laws, a regulatory and supervisory vacuum, and suboptimal coordination between authorities.

While frameworks may exist in a national context, given the crossborder nature of online shopping and cloud services, enforcement may be complicated. Given the international flow of data, it is likely that enforcement regimes and customer-recourse systems will not be clear, particularly in developing countries and in cases where data is held in the cloud and/or is unstructured data.

It important to establish adequate international frameworks and cooperation structures to address such issues. Relevant questions that emerge when looking at these issues might be: Do local data-protection laws deal effectively with these issues? What international frameworks and mechanisms could be set up? Is there a privacy/financial-sector enforcement authority that can address such issues? Does that authority effectively coordinate with other regulators, such as financial services authorities and telecommunications regulators, both at the national and international level?

BOX 3

Recently Issued Regulations and Guidance That Address Concerns about the Usage of New Types of Data from Multiple Sources

While legislation continues to be a key response to privacy risks, the focus on issues associated with the usage and processing of new forms of data, in particular those concerning big-data analytics, is limited. More and more jurisdictions have data-protection laws in place, and over 100 jurisdictions (around 50 of which are European) have adopted data-privacy laws. Nevertheless, these "new frameworks" are shaped around existing guidance and principles, discussed above, which, as analyzed, do not fully cover the emerging issues discussed in this note.

Despite this, in recent years some legislative, regulatory, and other initiatives have begun focusing more specifically on issues relevant to the usage of new types of data coming from multiple sources.

- **Examples of international initiatives.** Between 2012 and 2015, several national cybersecurity strategies were launched.[90] In addition, APEC has begun to review its 2004 privacy framework, and the International Conference of Data Protection and Privacy Commissioners has included "advancing global privacy in a digital age" among its strategic priorities for 2016–2018.[91]
- **The EU General Data Protection Regulation.** One important aspect is that the regulation will apply to all companies that target EU markets or consumers, broadening the range of controllers falling under its purview. Given the rise of geolocalization applications, location and other types of online identifiers have been included in the definition of personal data,[92] and restrictions on the processing of sensitive data have been expanded to include biometric identifiers.[93] Other key points in the regulation include (i) the "right to be forgotten," meaning that individuals have the right to request that their data be deleted when they no longer want it to be processed (subject to certain exceptions);[94] (ii) "the right to data portability" (for example, between financial services providers); (iii) "the right to know when one's data has been breached," addressing concerns of cybersecurity (this means that controllers will need to notify national supervisory authorities of data breaches, and data subjects will need to be notified of high-risk breaches); and (iv) data protection by design and by default. (See box 1.)
- **Consumer Protection Principles: Consumer-Authorized Financial Data Sharing and Aggregation.** On October 18, 2017, the Consumer Financial Protection Bureau issued a new set of principles that are not binding but are meant to provide guidance to a wide variety of stakeholders on specific issues relating to data protection and big data. In fact, the bureau recognizes that several stakeholders are working on ways to access, aggregate, and use customers' data, but it "believes that consumer interests must be the priority of all stakeholders."[95] Hence, it issued the principles to explain its vision of how the data-aggregation market can develop while also ensuring that customers are protected. Key issues covered by the principles include "(i) data scope and usability, (ii) control and informed consent, (iii) security, (iv) access transparency, (v) accuracy, and (vi) the ability to dispute and resolve unauthorized access."[96]

[90] For details, see *OECD Digital Economy Outlook 2017* (OECD, 2017), 225 and followings.

[91] 37th International Conference of Data Protection and Privacy Commissioners, "Resolution on Conference's Strategic Direction" (ICDPPC, October 27, 2015), available at **https://icdppc.org/wp-content/uploads/2015/02/Resolution-on-Conferences-Strategic-Direction-2016-18.pdf**

[92] Regulation (EU) 2016/679, Article 4(1).

[93] Regulation (EU) 2016/679, Article 9.

[94] Presumably, this provision is based on the "Right to Be Forgotten" decision of the European Court of Justice (C-131/12). For a summary, see, generally, **http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_data_protection_en.pdf**

[95] "Consumer Protection Principles: Consumer-Authorized Financial Data Sharing and Aggregation" (Consumer Financial Protection Bureau, October 2017).

[96] "Consumer Protection Principles."

Source: Own elaboration based on cited different sources.

7 CONCLUSION

The usage of new types of data coming from multiple sources has many benefits but also consumer risks, which need attention. It provides opportunities to expand access to financial services for financially excluded and underserved groups, to deliver better-suited, better-tailored products for consumers, and, ultimately, to reduce costs for providers, producing savings that can be passed on to consumers. However, there are also serious consumer concerns to be considered and further researched. They include a broken consent model and potential technology solutions, price and market-segmentation practices that are potentially discriminatory, ARCO rights that are difficult to implement and enforce, and the need to clarify acceptable security protocols.

There is a clear need for more detailed examination of the implications for financial consumers, with a view to developing an appropriate industry and regulatory response. There should be a focus on deep empirical and analytical research to determine the actual harms and possible industry, regulatory, and supervisory solutions to the identified issues. Any such research should involve wide consultation with a broad range of stakeholders, including industry, regulators, consumer groups, academics, and international development agencies. On the regulatory front, this should include not just consumer and data-protection regulators; financial-sector, telecommunications, and competition- and business-development agencies should also be involved in the discussion.

There should also be close coordination between all relevant international forums working on these issues. The aim should be to monitor, oversee, and share information about relevant issues and possible solutions, with a view to ensuring that consumers are adequately protected while safeguarding the benefits arising from innovations. The World Bank recognizes that big data is an emerging issue and provides guidance in "Retail Payment Services," annex A of its recently issued 2017 edition of the Good Practices, as to how good practices relating to data protection and privacy can be applied in a big-data context.[97]

[97] See "Retail Payment Services," annex A of *Good Practices for Financial Consumer Protection*, 2017 Edition (World Bank Group, 2017), Good Practice D1, explanatory notes.

WORLD BANK GROUP