2018

Data Protection and Privacy for Alternative Data

GPFI- FCPL SUB-GROUP DISCUSSION PAPER -DRAFT- MAY,4 2018

THE WORLD BANK AND CGAP

Table of Contents

ACKNOWLEDGEMENTS

This document has been prepared by the World Bank and the Consultative Group to Assist the Poor (CGAP) based on the terms of reference established by the G20 GPFI Presidency for 2018. Its development has taken into consideration previous work conducted by the GPFI in the matter (e.g., High Level Principles on Digital Financial Services, the current work being conducted by the GPFI- Sub Group on SME Finance, the work developed by the World Bank on the matter (e.g., Good Practices on Financial Consumer Protection, country surveys and studies on the matter) as well as existing international standards and regional legal and policy frameworks (e.g., The General Principles for Credit Reporting, The EU General Data Protection Regulation (GDPR), the Asia Pacific Economic Cooperation (APEC) Privacy Framework and the OECD Guidelines on Privacy).

An outline and presentation of the content of the guidance was also shared with members of the Global Partnership on Financial Inclusion (GPFI) ahead of their meeting on 7 February 2018 in Buenos Aires, for comment through the written process by 19 February 2018.

This document has benefited from a consultation process with members of the respective sub-groups of the GPFI namely the Subgroup on SME Finance and the Subgroup on Consumer Protection and Financial Literacy. In addition, the document also incorporates many of the findings included in the World Bank report “New Forms of Data Processing Beyond Credit Reporting: Consumer and Privacy Aspects”, 2018, and the Responsible Finance Forum report “Opportunities and Risks in Digital Financial Services: Protecting Consumer Data and Privacy” 2017.

This work was led by Margaret Miller (Lead Financial Sector Economist, The World Bank) and main authors include David Medine (Senior Financial Sector Specialist, CGAP) and Fredes Montes (Senior Financial Sector Specialist, The World Bank).

ACRONYMS

AML	Anti-Money- Laundering
APEC	Asia Pacific Economic Cooperation
ARCO	Access, Rectification, Cancelation and Opposition
CFT	Combating the Financing of Terrorism
CGAP	Consultative Group to Assist the Poor
CRSPs	Credit Reporting Service Providers
GATS	General Agreement on Trade in Services
GPCR	General Principles for Credit Reporting
GDPR	General Data Protection Regulation
HLP-DPFI	High Level Principles for Digital Financial Inclusion
IADB	Inter-American Development Bank
ICCR	International Committee on Credit Reporting
ICT	Information Communication Technologies
IMS	Identification Management System
MSMEs	Micro Small and Medium Enterprises
OECD	Organization for Economic Cooperation and Development
SMEs	Small and Medium Enterprises

I. Introduction and Background

1. **The financial services industry is undergoing a fundamental transformation due to digital technology: digitization both facilitates access to finance and increases the scope for efficient and cost-effective collection and analysis of enormous quantities of data for use in financial decision-making.** Data used to predict credit risk can now go beyond customers' payment history on other formal loans to include their payment behavior with other kinds of businesses and utilities, their personal spending habits, and other traceable information related to use of e-commerce sites, apps and geolocation. Personal data that may indicate both ability and willingness to repay may also be gleaned from social media and from psychometric evaluations where data are collected through surveys or loan applications.

2. **Lack of data has traditionally been an obstacle to financial inclusion for many people in developing countries, especially for those who operate in the informal sector.** People in the informal sector typically include the poor living in urban areas, as well as most migrants, many rural inhabitants, women and young adults. These groups often have low and/or irregular incomes, and many of them are also informally employed. Much of the economic and financial activity of unserved and underserved individuals takes place in the informal sector, where transactions are not recorded. Moreover, in some societies women and/or undocumented migrants, among others, may not be able to have accounts or credits under their own name, and as a result they cannot build a personal credit history. The lack of credit history, known as "thin files" by credit bureaus, impedes access to credit from banks, financial institutions or even from real sector formal creditors.

3. **Expansion of the use of alternative data can bring individuals and MSMEs currently operating in the informal sector into the formal sector.** While there are many reasons why some might operate informally, removing a constraint on entry into the formal sector will encourage some and permit others who are motivated to shift their financial activity into the formal sector. As noted, currently a major obstacle to entry into the formal sector is the lack of a traditional credit history. Alternative data can reduce or remove that barrier by providing risk assessment tools to formal financial institutions.

4. **In an increasingly digitized world, vast quantities of "alternative data" are being generated every day which can complement or substitute for traditional financial data (such as information on loan payments, defaults and bankruptcies) and open the door to financial services for previously unserved or underserved customers.** In some instances, alternative data are being created outside the financial system, such as through e-commerce platforms. These data – on commercial transactions, cash flow, client base, even customer reviews – can then be used by non-traditional financial services providers / fintech companies to evaluate risk, determine credit capacity and to offer financial services to firms or individuals who may lack access to bank credit. Alternative data are also useful for traditional lenders; banks may seek to leverage alternative data, such as information from utilities or retail lending, to reach new customer segments including micro

and small enterprises.[1] Beyond being used to provide access to credit, alternative data may offer valuable granularity on customer preferences and behaviors that can be used to design new financial products and services, encourage positive financial behaviors and support the real sector by linking financing to energy, commerce, health or other sectors.

5. **While offering many benefits, the use of new types of alternative customer data for financial and other sensitive decisions also raises significant data protection and privacy concerns including revealing confidential information to third parties, enabling aggressive marketing practices, and security risks including fraud and identity theft**. A key component of data protection is data security. With a greater reliance on electronic communications and interconnected transactions comes the risk that cybercriminals could hack into information systems, disrupting them and potentially stealing data. This increased vulnerability to fraud and disruption could negatively impact access to financial services in the future.

6. **Policy makers face the challenge of striking the right balance between promoting the benefits of the expanded use of alternative data while ensuring adequate data protection and attention to consumer privacy across the eco-system.** This is a challenge made still more difficult by the fast pace of technological change, limited knowledge of the actual risks and their consequences for consumers in cyberspace, and fragmented (at best) legal and regulatory data protection frameworks covering financial services.

7. **Data protection and privacy requires a new way of thinking and preparation as regulatory or institutional frameworks to protect individuals and firms either do not exist or are being rapidly outpaced by technological advances.** Shared information infrastructure can harness data in ways that provide a more complete picture of the customer to financial services providers. At the same time, shifting from a narrower set of data providers and users, as has often been the case with bank-centric credit bureaus, to a very broad and even undefined set of data sources and users creates new challenges in terms of data control, data accuracy, storage, protection and usage. Done right, legislation can promote consumer welfare, encourage economic development, respect human rights and enhance competition. A further challenge will be the use of algorithms that blend traditional and alternative data in ways that are complex and not always transparent. The risks associated with data-driven business models, and particularly data protection and cybersecurity, should not be forgotten.

8. **Inadequate data protection practices, standards and rules can result in consumers experiencing financial harm, loss of privacy and reduced trust**. In the absence of international standards on data protection, there are several internationally agreed-upon frameworks that provide guidance on protection of personal data and privacy. Key frameworks on data protection include the European Council,[2] the Organization for Economic Co-operation and Development

[1] For further information on predictive nature of credit scores using alternative data please see "Predictive Value of Credit Scores" Center for Financial Innovation Services, 2018.

[2] The European Council comprises 47 members and issued in 1981 the Treaty No. 108, *Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data,* which was ratified by 51 countries. https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/108.

(OECD),[3] the International Conference on Data Protection and Privacy Commissioners (ICDPPC),[4] and Asia-Pacific Economic Cooperation (APEC).[5] It is also relevant to mention the European Union (EU) General Data Protection Regulation (GDPR) came into force May 2018 and applies to 27 countries plus includes the extraterritorial concept which might have an impact in other jurisdictions beyond EU. Finally, the Directive on Privacy and Electronic Communications[6] (e-Privacy Directive) also concerns electronic communications, confidentiality and privacy and complements the GDPR. While data protection legislation needs to be tailored to the specific needs of the country enacting it, each country must take into account various business models in its economy, cross-border trade, state of technological advancement, social and political values and legal systems.

9. **Under Argentina's Presidency of the G20 and coordination of the GPFI, work on data protection and privacy focuses on the development of high-level guidance for policymakers on opportunities and challenges related to the beneficial use of alternative data including policy options.** Against this background, the objective of this paper is to provide guidance for countries to use in considering key options to mitigate risks associated with data protection and privacy, especially related to alternative data, but not to direct adoption of unduly prescriptive legislative provisions. The rest of the document is organized as follows: Section II includes a discussion on the definition of alternative data and scope of the paper; Section III illustrates the key data protection and privacy considerations related to the use of alternative data and provides policy guidance; Section IV presents issues of coordination amongst authorities and discusses potential solutions to enhance such coordination.

[3] The OECD issued *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data,* in 1980 which were revised in 2013. http://www.oecd.org/sti/ieconomy/privacy.htm.

[4] Notably the *Madrid Resolution* adopted in 2009. https://edps.europa.eu/sites/edp/files/publication/09-11-05_madrid_int_standards_en.pdf.

[6] Directive on Privacy and Electronic Communications, irective 2002/58/EC and the 2009 update, Directive 2009/136 concerns electronic communications and the right of confidentiality, data/privacy protection and more.

1. POLICY GUIDANCE: Lawful Collection of information

Alternative data involving personal information used for the evaluation of creditworthiness of consumers and SMEs should be collected and processed lawfully. The legal bases could involve consumers' consent for collection and processing that is necessary; (i) for the performance of a contract to which the data subject is party, (ii) for compliance with legal obligations, (iii) to protect vital interests of the data subject, (iv) for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller and (v) for the purposes of the legitimate interests pursued by the controller or by a third party.

2. *POLICY GUIDANCE: Transparency*

Industry participants should enable mechanisms that allow consumers to understand the key facts (e.g., name of the data controller, purpose of the data collection, potential users of the data, consumers' rights, details on dispute handling mechanism and lawful bases for such data collection) of the data collection. Informing consumers about these facts, contributes to enhanced transparency and trust. This could be achieved in the form of a privacy policy or privacy notes which could be provided through electronic means or any other ways that are convenient to consumers in an easy to understand language and avoiding lengthy legalistic wording.

3. *POLICY GUIDANCE: Accuracy and reliability*

Regulators, policy makers and industry participants should adopt measures to ensure that alternative data collected is lawful, matched to the correct person, obtained from trustable sources, up-to date and relevant to the purpose for what is being used. Understanding that data will not be completely free of error, accuracy of data obtained from multiple sources to make risk evaluation decisions could be achieved by establishing requirements relating to the collection, data processing and further distribution of the information.

Box 1- Summary of Policy Guidance on Data Collection

Box 2- Summary of Policy Guidance on Processing and Further use of Data

4. POLICY GUIDANCE: Accountability

Industry participants using alternative data for creditworthiness evaluation should be able to identify the data flows path from the original source. Data controllers should be responsible for taking adequate measures to implement data protection and privacy principles and be able to demonstrate that they have taken appropriate measures to ensure protection of personal data they are responsible for from origin to use.

5. POLICY GUIDANCE: Consent

Recognizing that there are other legal bases for data collection and further processing, when alternative data is being used for a different purpose than the one specified for the data collection -consistent with domestic laws- consumers' consent shall be necessary. This policy guidance is also applicable for cross-border data flows. When data is provided by a third party, cost efficient mechanisms should be in place to enable consent where required.

6. POLICY GUIDANCE: Consumer's Rights

Data controllers should enable mechanisms that allow consumers access and the ability to correct their information as well as request the deletion of data -when appropriate- based on applicable laws or rules on retention periods. In addition, consumers should also be able to object to the processing of their information for certain purposes (i.e., marketing). Consumers should also be given the opportunity to transfer their data to any other service provider of their choice without affecting the usability of the data.

7. POLICY GUIDANCE: Security

Consistent with existent national cybersecurity plans, industry participants should conduct periodic cybersecurity risk assessments, develop policies and procedures to effectively respond to cyber incidents, communicate cyber incidents to all relevant parties -including consumers- as soon as practicably possible or as required by law, and devote resources to assess, monitor and mitigate consequences of cyber-incidents. These measures are also applicable for any outsourced service that involves the processing and storage of personal information.

Box 3- Policy Guidance on Fair Discrimination and Cross-Border Data Flows

8. POLICY GUIDANCE: Discrimination

Policy makers and industry participants should adopt measures to ensure that predictiveness of alternative data is tested and verified, that scoring models developed using alternative data do not unfairly discriminate against protected groups. The use of alternative data that carries forward historical discrimination is either prohibited or restricted, taking into account its ability to predict risk and the availability of alternative decision-making tools;

9. POLICY GUIDANCE: Cross-Border Data Flows

Authorities should enable cross-border data flows when appropriate and consistent with privacy laws and international privacy frameworks. Cross-border data flows should be subject to mechanisms that ensure accountability of data controllers and industry participants. There should be in place procedures and policies to allow consumers to implement their rights - - regardless of where the data is stored or has been transferred. Finally, cooperation agreements between authorities -when a legal framework is in place- allows for adequate implementation of privacy rules across borders.

10. POLICY GUIDANCE: Cooperation and Coordination

Authorities need to develop a National Strategy towards privacy and set up coordination mechanisms between authorities and industry participants at domestic and international level. Ideally, a legal framework on data protection and privacy inspired in internationally agreed-upon framework should be established. Identification of roles (e.g. enforcement, rulemaking and supervision) regarding data protection and privacy of each authority as well as the establishment of formal cooperation mechanisms would allow for smooth implementation. Dialogues should be fostered between financial sector authorities, ICT, consumer protection and data protection and privacy authorities for the appropriate implementation of privacy rules related to the collection, processing and further use of alternative data.

II. Definition of Alternative Data and Paper Scope

10. **The scope of the paper will include attention to both individual consumers, microenterprises / sole-person firms and to small and medium-sized enterprises (SMEs).** There is no globally accepted definition of SMEs and therefore the concept of SME will be defined at country level. For small firms, including microenterprises and sole proprietorships or sole-person firms, it is often the owner's personal credit history and financial capacity which is used as the basis for granting credit and risk analysis. Data protection and privacy as these relate to alternative data for these smaller firms may be overlapping, or similar to, individual protections. Also for the clarity of this background paper, it should be acknowledged that data protection frameworks often do not cover legal entities under their scope of protection unless such information refers to individuals as regards to their personal sphere as opposed to their professional sphere.

11. **There is no single, widely accepted definition for alternative data, but there is a clear common approach to this topic, which emphasizes access to enormous quantities of data created through digital means.** For example, the GPFI Priorities Paper 2018 defines alternative data as, "A generic term that designates massive volume of data that is generated by the increasing use of digital tools and information systems." The ICCR defines alternative data as "information readily available in digitized form that is collected through technological platforms[7]" and traditional data as "individual's history of fulfilment of his/her financial and other similar obligations"[8]. The GPFI included in the report on "*Use of Alternative Data to Enhance Credit Reporting to Harness Digital Financial Services*" a definition of two broad categories of alternative data namely; (i) structured data "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" and (ii) unstructured data referred to as "information that either does not have a pre-defined data model and/or is not organized in a predefined manner". Such report recognizes that alternative data might have different meanings depending on the jurisdiction and therefore in the context of credit reporting service providers (CRSPs), the definition of alternative data will remain country specific and dependent on the kind of credit information that CRSPs in these jurisdictions are currently collecting. As a result, what is alternative in one market can be traditional in another.

12. **Relevant information to evaluate consumers' risk has evolved over time along with the sources of such data.** In traditional credit reporting, information collected to evaluate consumers' creditworthiness included: (i) identification and geodemographic data, (ii) credit borrowing and repayment data, (iii) additional information related to secured interests, (iv) information on court judgements and tax liabilities and (v) available information from government held databases mostly from companies' registries or similar sources. Data providers were often

[7] Please see background document on "Use of Alternative Data to Enhance Credit Reporting to Harness Digital Financial Services to Individuals and SMEs operating in the Informal Economy", ICCR, March, 2018
[8] "Credit Reporting Systems Contribution to Financial Inclusion", The World Bank, 2017 (page 5)

limited to formal financial institutions or clearly recognized creditors such as retailers that provided store credit, and the data that was shared emphasized repayment history as a way to promote positive credit behaviors and facilitate risk analysis across the financial market. This information was complemented by publicly available information, such as data from courts and other government databases.

13. **Early efforts to use "alternative data" in mainstream credit markets were in relation to consumers with so-called "thin files" who lacked previous banking relationships: alternative data sources included payment data related to non-loan products requiring regular payments such as rent, insurance, utilities or telecommunication services bills.** Consumer subjects of alternative data were most likely to be outside the economic mainstream in some way, such as minorities, women, migrants or young adults. Some consumers did not have any previous credit history for several reasons including never having applied for credit, migrants (perhaps with a credit history elsewhere), or being customers of financial providers such as microfinance institutions which were not part of the credit reporting system. To evaluate the risk these "thin file" customers present, additional sources were included in many credit reporting systems. For instance, data on utility bill payment records were considered as relevant information to evaluate repayment behavior.

14. **More recently the introduction of digital services, new payment channels, adoption of digital technologies and smart devices have enabled the collection of data on consumers' behavior patterns regardless of whether the data referred to any credit repayment or payment for goods or services.** New sources of data include electronic commerce, on-line shopping, application-based services, transaction data and cash flow, residential stability, educational attainment, occupational status, daily routines and social networks. Data collected through mobile phones and telecommunications (e.g., call data records, airtime top ups, P2P, G2P and P2G payment transactions) are also exponentially increasing data trails including for low income consumers in developing and emerging markets[9].

15. **For the evaluation of SMEs creditworthiness, the GPFI report on Alternative Data includes two categories: one related to data that can be directly used as input in formulas for statistical and analytical tools and other data which is more abstract and requires subjective analysis**. The first category of data refers to cash flow, sales volume, shipping data, merchant transactions, and mobile money transactions among others. The second category of data which requires subjective analysis to determine its relevance to the expected outcome includes psychographic attitudes toward credit, mobile usage (e.g., number of calls to the same number, calls made on peak usage, etc.). This information refers to attitudes, abilities, personality traits and is gathered through questionnaires and surveys, similar to the pre-employment testing and software tools used for a long time by recruiting or talent assessment companies. This type of information could also include customers' reviews of their suppliers' business, how SMEs interact with their customers; which are the result of personal behaviors or attitudes and which are available through social media.

[9] Some specific work has been produced as a follow-up to this initiative, including the joint report of the Committee on Payments and Market Infrastructures (CPMI) and the WBG on the "Payment aspects of financial inclusion" released in April 2016 (http://www.bis.org/cpmi/publ/d144.pdf).

16. **Personal Data refers to the information of an identified or identifiable natural person.** "any information relating to an identified or identifiable natural person (data subject), an identifiable natural person is one who can be identified directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic mental, economic, cultural or social identity of that natural person". Information that has been anonymized is not typically subject to data protection rules but technology allows the re-identification of a person using information available and anonymized data. To this extent, the data protection and privacy aspects in relation to SMEs relate to those situations where information about an individual (e.g., owner or management) is used to analyze the risk of the SME. This information could be in the form of a report or embedded as one of the attributes of the score.[10]

17. **This paper takes into account the many new sources of data available through digital technology, and which now can constitute "alternative data" for credit granting, financial risk assessment, collection and business development purposes.** The specific sources and types of data which may be included among alternative data for the purposes of this paper include: (i) transaction data (this includes more granular information from payment services, such as credit cards as well as other digitally tracked transactions such as through e-commerce); (ii) telecommunications, rent and utility data; (iii) social profile/social media data; (iv) audio and text data (often collected through recorded customer service, application or collection calls and outreach); (v) app and clickstream data (collected as a customer uses an app or, for clickstreams, moves through a website); (vi) social network analysis (can include compiling a comprehensive data file for a customer across providers and data sources and/or analyzing the people with whom a customer is connected); (vii) Internet of Things (IoT) which includes data from smart grids, smart devices and shipping and transport systems; (viii) crowdsourced data such as reviews from online communities and specialized social networks; (ix) weather and satellite data; and (x) survey and questionnaire data including psychometrics.

18.

.

19. **The paper will also include data protection and privacy considerations for alternative data as they may disparately affect individuals and firms in the informal sector where sources of formal data are scarce and where the ability of consumers to complain may be limited by their informal status.** Informality here covers situations where economic actors (individuals or legal entities / firms) are not part of the formal economy due to a lack of general compliance with laws, regulations, fiscal commitments and reporting and other required public fees and registrations.

---

the *Madrid Resolution* adopted in 2009. https://edps.europa.eu/sites/edp/files/publication/09-11-05_madrid_int_standards_en.pdf.

[10] Full APEC framework please see"https://cbprs.blob.core.windows.net/files/APEC%20Privacy%20Framework.pdf" https://cbprs.blob.core.windows.net/files/APEC%20Privacy%20Framework.pd.

III. Identified Risks to be Addressed Through Policy Guidance

20. **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: data used to conduct risk assessment is inaccurate or non-reliable; the use of personal information is based on uninformed and meaningless consumer consent; illegal discrimination; unfair price segmentation; lack of transparency about the collection, use, and disclosure of personal information; insufficient data security; and failure to provide effective access and correction and complaint-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. To illustrate some of the risks faced by regulators under this new scenario, the figure below shows the perspective of the Banco de Mexico on the use of alternative data.

Figure 1- Major Risks perceived by Authorities in Mexico on the use of Alternative Data

1. The alternative data providers (different from credit bureaus) might not ensure data quality (inconsistent, incomplete or inaccurate);
2. The information could come from non reliable sources;
3. Infringements of consumer privacy (sensitive data) or discriminatory issues could arise.
4. Alternative data providers might not have rigorous security controls or standards;
5. Lack of transparency and lack of the possibility to correct data as, for instance, the data subject might not have effective judicial or extrajudicial dispute resolution mechanisms against alternative data providers;
6. Consumer cannot change their behavior in order to improve their credit rating, and
7. Entities could take the information obtained from alternative data providers as a substitute of the credit reports given by credit bureaus.

Alternative data providers in Mexico have to comply with the Data Protection Law.

Source; Banco de Mexico, March 2017

III.1. Data Collection

POLICY GUIDANCE: Lawful Collection of information

Alternative data involving personal information used for the evaluation of creditworthiness of consumers and SMEs should be collected and processed lawfully. The legal bases could involve consumers' consent for collection and processing that is necessary; (i) for the performance of a contract to which the data subject is party, (ii) for compliance with legal obligations, (iii) to protect vital interests of the data subject, (iv) for the performance of a task carried out in the public interest or in the exercise of official

authority vested in the controller and (v) for the purposes of the legitimate interests pursued by the controller or by a third party.

Key Considerations

- ***Consider minimization of data collection**. Regulators could identify key data items that are relevant for risk evaluation, and only those data items that should be captured and used under specified circumstances or allow industry participants to demonstrate the relevancy of such data to the purpose of risk evaluation. This concept envisions that only the minimal amount of data should be collected. The GDPR covers this principle under its Article 5(1)(c): "Personal data shall be… adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimization')." In the case of alternative data, structured information related to credit repayment behavior -including positive information- should be considered relevant as well as identification information. Beyond that, it is clear that not all types of information collected qualify as relevant.*
- ***Authorities should define "legitimate interest" at a domestic level**. The concept of legitimate interest might vary country to country. In general, the use of credit repayment behavior to evaluate risk might be considered legitimate interest when shared with CRSPs but additional non-structured data requires further justification.*
- ***Publicly available data** -When collecting information available in the public domain from government held databases, it might not be feasible or appropriate to obtain consent, although rules regarding accountability, data accuracy and consumers' rights might still apply. If information obtained from the public domain is being used for a different purpose than the one established in the data collection, consent would be necessary.*
- ***Compliance with legal obligations**. There are clear exemptions in the financial services world to the need for consent (i.e., use of personal information to evaluate suspicious transactions related to AML or CFT, data collection for the evaluation of credit risk under a regulated networked environment).*

POLICY GUIDANCE: Transparency

Industry participants should enable mechanisms that allow consumers to understand the key facts (e.g., name of the data controller, purpose of the data collection, potential users of the data, consumers' rights, details on dispute handling mechanism and lawful bases for such data collection) of the data collection. Informing consumers about these facts, contributes to enhanced transparency and trust. This could be achieved in the form of a privacy policy or privacy notes which could be provided through electronic means or any other ways that are convenient to consumers in an easy to understand language and avoiding lengthy legalistic wording.

Key Considerations

- ***Adoption of Privacy Notices**. Regulators could promote transparency on service providers by suggesting the adoption of privacy notices. These notices could be embedded in the design of new digital products and services and could be provided to consumers at the moment of the data collection. Privacy notices should be simple, concise, easy to adopt; regulators should aim to provide flexibility to industry participants' finding ways that are proportionate and efficient.*

- ***Disclosure of information on scoring models**. When information collected is used to develop a scoring model, consumers should be informed about the key attributes that are included in the development of such scores. In addition, the key facts of the scoring model should be disclosed to consumers as well as the sources of the information used for the scores.*
- ***Adverse action**. When using alternative data for automated decision making, consumers should be informed of any adverse action taken against them based on such data, as well as the key characteristics that led to such decision. In this case, the users of the final product (score, report or other informational product) should find ways to effectively notify consumers regarding such actions and the key attributes that led to such decision.*

POLICY GUIDANCE: Accuracy and reliability

Regulators, policy makers and industry participants should adopt measures to ensure that alternative data collected is lawful, matched to the correct person, obtained from trustable sources, up-to date and relevant to the purpose for what is being used. Understanding that data will not be completely free of error, accuracy of data obtained from multiple sources to make risk evaluation decisions could be achieved by establishing requirements relating to the collection, data processing and further distribution of the information.

Key Considerations;

- ***Industry participants[11] could adopt a privacy by design approach**. 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. 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.[12]*
- ***Effective identification** of the consumer enhances data accuracy. By the enhancement of Identification Management Systems (IMS) and adoption of digital identities, information is more likely to be linked to the proper consumer. Industry participants should deploy methods to verify the identification of consumers using IMS. For data to be accurate, it is important that rigorous steps be taken to match data to the correct person which might, among other things, require proper identification of individuals and legal entities.*
- ***Information to be collected from trustable sources**; for example, entities that have a strong internal reason to maintain accurate, complete and timely records.*
- ***Data is up to date** with frequency of updates being set by the importance of timely information and technological limitations, typically monthly or shorter intervals. When unstructured data is collected*

[11] Industry participants refers to banks, financial institutions, Fintechs, Mobile Network Operators and similar entities.

[12] 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; David Medine, "Privacy by Design for Financial Services" https://www.livemint.com/Opinion/1ShpKAOC59VlXiwgCkVv8O/Privacy-by-design-for-financial-services.html.

from different sources, industry participants acting as data controllers should ensure that data used in the evaluation of creditworthiness is updated on a systematic basis and such changes are reflected in the algorithms built with such data.

- ***Data collected and used meets evidence accuracy tests**; At the minimum the adoption of validation and normalization rules to consider alternative data for risk evaluation taking into account how predictive is the data for future behavior should be fostered.*[13]

Explanatory Text on Data Collection

21. **The GDPR provides guidance on the interpretation of "legitimate interest" in the context of EU Member States**. Article 29 Working Party states that "*The legitimate interests of a controller, including those of a controller to which the personal data may be disclosed, or of a third party, may provide a legal basis for processing, provided that the interests or the fundamental rights and freedoms of the data subject are not overriding, taking into consideration the reasonable expectations of data subjects based on their relationship with the controller*". Such legitimate interest could exist for example where there is a relevant and appropriate relationship between the data subject and the controller in situations such as where the data subject is a client or in the service of the controller. As an example, collecting personal data to avoid fraud constitutes legitimate interest and so does the collection of data for network security, information on credit card payment performance to be shared with credit reporting service providers (CRSPs) constitutes legitimate interest based on interpretation from the UK Information Commissioner Office (ICO).

22. **Informing consumers about the collection and processing of data contributes to increased consumers trust**. Right to be informed about the collection, processing and further disclosure of information is recognized under the APEC Privacy Framework: "Personal information controllers should provide clear and easily accessible statements about their practices and policies with respect to personal information."[14]

23. **Despite the notion of Big Data which involves the collection of as many data as possible, when collecting alternative data to be used for creditworthiness evaluation consideration should be given to what constitutes relevant data for such purpose**. The General Principles for Credit Reporting (GPCR), under GP1, establish that "data collected should include all relevant information to enable any given user to adequately evaluate and manage credit risks on a continuous basis." The GPCR establishes a limit on the data that can be shared which is associated with the permissible purposes underlying information sharing or privacy considerations when dealing with sensitive issues such as ethno-demographic data. In addition, the ICCR suggests the adoption of a data protection framework that guards disadvantaged individuals' data by limiting its uses and access considering also the range of data uses which can benefit individuals. The

[13] Katy Jacob, "Reaching Deeper: Using Alternative Data Sources to Increase the Efficacy of Credit Scoring," CFSI, March 2006.

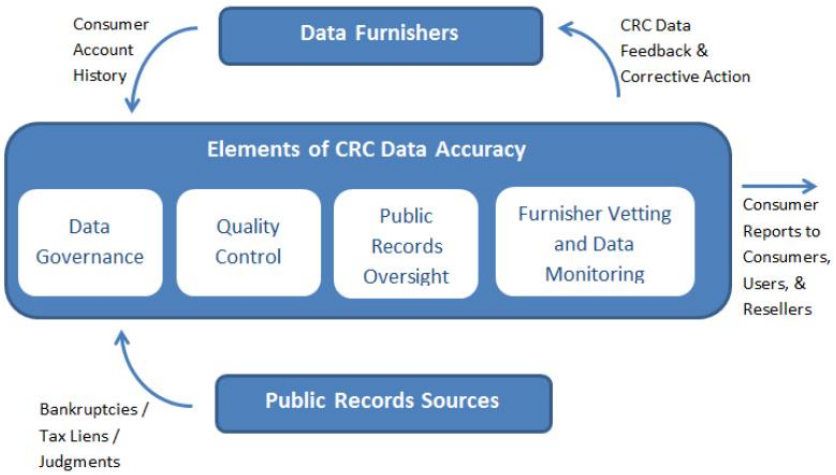
[14] Please see Principle on Notice which includes also a list of minimum informational aspects; (i) personal data being collected, (ii) purposes for which such data is being collected, (iii) types of organizations to whom personal data might be disclosed, (iv) identity and location of the data controller and (v) choices that data controller offers individuals on the use further disclosure and dispute handling.

specific sensitivities in each country when it comes to privacy issues to be considered when putting together this framework, although a minimum general understanding on these issues should be sought in order not to hamper cross-border activities.[15]

24. **Internationally agreed-upon frameworks on data protection and privacy recognize several legal grounds that justify the collection of personal information**. For data collection to be lawful it should be based on the accepted existing grounds for data collection. While in some cases consent may be one of them, it is not the only one. Other bases include: (i) need for such processing for the performance of a contract to which the data subject is party or (ii) processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party.

25. **For data to be meaningful and useful it has to be reasonably free of error, truthful, complete and up to date. In addition, for data to be reliable it should meet evidence tests**. The Principle 5 of the High Level Principles for Digital Financial Inclusion (HLP-DFI) calls for the development of guidance to ensure the accuracy and security of all data related to accounts and transactions, digital financial services and development of credit scores for financially excluded and underserved consumers.[16] The OECD Data Quality Principles (Principle 8) provide 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). In credit reporting systems, GP1 on Data requires that data collected and distributed should be -to the extent possible- free or error, truthful, complete and up to date. In a controlled environment (closed network) data accuracy is easier to detect, mitigate and respond to than in an open environment.

Figure 2-Example of data accuracy graph on credit reference agencies



[15] Please see "Policy Brief on Credit Reporting and Financial Inclusion" The World Bank, 2017.

Source: Inter-American Development Bank and World Bank Workshop "Credit Reporting in the Americas", March 2017

26. **Accuracy problems emerge when small and large 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 could be questioned when data is initially captured for one purpose and from one source and is then used for a very different purpose and by complete different users.
- Sources of data might not always be trusted.
- Data does not meet evidence tests and when challenged is proven to be erroneous or false.
- Data not up-to date.
- Data association might not be subject to robust tools and might lead to incorrect inferences about the consumer.[17]
- Data sources might not be active data providers or for other business reasons might not have sufficient incentives to maintain data up to date, might not collect data in similar formats, or check for the veracity of reported data.

III.2 Processing and further use of data

POLICY GUIDANCE: Accountability

Industry participants using alternative data for creditworthiness evaluation should be able to identify the data flows path from the original source. Data controllers should be responsible for taking adequate measures to implement data protection and privacy principles and be able to demonstrate that they have taken appropriate measures to ensure protection of personal data they are responsible for from origin to use.

Key Considerations

- **Clear data flows path**- Identification of data source is key to identify all the relevant parties in the data flow path.
- **Keeping records of data disclosure**- mechanisms should enable the identification of data disclosures between data controllers, processors and users.

POLICY GUIDANCE: Consent

Recognizing that there are other legal bases for data collection and further processing, when alternative data is being used for a different purpose than the one specified for the data collection -consistent with domestic laws- consumers' consent shall be necessary. This policy guidance is also applicable for cross-border data flows. When data is provided by a third party, cost efficient mechanisms should be in place to enable consent where required.

[17] See Katy Jacob and Rachel Schneider, "Market Interest in Alternative Data Sources and Credit Scoring," The Center for Financial Services Innovation, December 2006.

Key Considerations

- ***Tiered consent**. 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 collection and processing of certain types of data or for specific purposes. Adopting a consent model that enables consumers to decide the type of data that they choose to share and the service providers that they allow to access their information. In adopting this measure, regulators should bear in mind that there are certain circumstances and data items that do not allow for consent (e.g., use of default data on credit repayment, or use for illegitimate purposes).*
- ***Consent to be given in any format convenient to the consumer**- Ideally, consent should be given electronically which might call for enhancement in systems to make it easy to collect and record consent methods. For example, through SMS in mobile devices. Likewise, revocation of consent should be made as easy as it is to provide consent.*
- ***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. In the case of traditional data used to evaluate risk (i.e., credit repayment data) this solution might not apply but instead rely on the date of expiration of contract to further process data unless additional consent has been provided by the consumer..*
- ***Opt-in, rather than opt-out, consent.** The recitals to the General Data Protection Regulation state that "silence, pre-ticked boxes or inactivity should not therefore constitute consent."[18] Industry participants should enable this feature by including clear processes to ensure that consumers receive all the relevant information to make their choice. Technological features and consent management systems could facilitate this process.*

- ***Record evidence of consent provided**. Industry participants should be responsible to record evidence of consent being collected from consumers. This is even more relevant when data is to be shared with third parties. For this process a consent management system would be useful.*

Explanatory Text on Processing and further use of data

27. **The number of different data sources, formats and collection methods involved in the use of alternative data (structured data and unstructured data) poses challenges to the adequate identification of responsibility**. Such complicated technological landscapes with many entities involved and difficult to track processing operations entail an inherent risk of errors and failures, while threatening the inability to identify the responsible party. Without a clear identification of data flows between the industry participants and their contribution to the ongoing processing operations, it is difficult to assess how the burden of accountability for data protection is distributed. Article 5 (2) of the GDPR includes the notion of accountability: "*The controller shall be responsible for and able to demonstrate compliance with …*". This notion of accountability is also recognized by the APEC framework.

[18] For further guidance, please 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 32.

28. **Consent is a fundamental principle concerning data privacy and Financial Consumer Protection**. Various standards, however, define and deal with consent differently. Principle 5 of the HLP-DFI encourages meaningful choices by consumers.[19] 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."[20] 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.[21] The GDPR calls for a "valid, freely given, specific, informed and active consent". 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).[22] The APEC Privacy Framework also refers to the need for consent to the collection of personal information (where appropriate). 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.[23] 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."

29. **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.

30. **Based on consumers' surveys and analysis conducted on consent clauses, a number of limitations to the current consent model have been identified.** Despite the common concern that current practices do not allow for real choice or real notice regarding the collection and further use of consumer data, a number of specific barriers to the adoption of an effective consent model were identified: (i) Consent clauses are typically tied to standard adhesion contracts impeding consumers to opt-out of any of the clauses or negotiate them in any way; (ii) consent clauses are extremely lengthy and difficult to understand; (iii) consent clauses do not include the necessary information to

[19] Ensure consumers of digital financial services have meaningful choice and control over their personal data—including through informed consent based on clear, simple, comprehensive, age-appropriate and brief privacy policy disclosures in relevant languages. Consumers also need to have transparent, affordable and convenient access and correction rights which can be exercised via remote and Internet-enabled access, including mobile phones and websites—or via a 24-hour call center.

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

[21] OECD Guidelines, Principle 10.

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

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

make an informed decision may not be provided to consumers;[24] (iv) hidden forms of consent as well as multiple consent clauses; (v) lack of choice as consumers do not have a meaningful choice if they want such product or service; (vi) the use of consumers' information is subject to opt-out rather than opt-in making it more difficult for consumers to implement their choice; (vii) evidence suggests that consumers rarely if ever read consent clauses; and (viii) finally, consumers are not always able to withdraw their consent once given.

31. **The Article 29 Working Party[25] Guidelines on Consent issued in April 2018 provide additional guidance on the implementation of consent as established under article 4 of the GDPR**. The guidelines recognized that since the adoption of the concept of consent under the Directive 96/45, then the introduction of the E-Privacy Directive (2002/58/EC) and now the GDPR, the notion of consent has evolved. However, under the GDPR conditions for consent are considered pre-conditions for lawful processing (article 6 (a) of the GDPR). Article 4 of the GDPR defines consent as "*any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her*". In addition, the GDPR allows data subjects to withdraw their consent (reversible decision of data subjects) at any time. The article 29 working party guidelines explain that:

(i) If a data subject feels compelled to consent or will endure negative consequences if they do not consent, then consent will not be valid;

(ii) If consent is bundled up as a non-negotiable part of terms and conditions it is presumed not to have been freely given;

(iii) consent will not be considered to be free if the data subject is unable to refuse or withdraw his or her consent without detriment.

Figure 3- Opt-in model of consent to capture data for credit risk evaluation

[24] Please see additional reading for informed consent. 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.

[25] Article 29 Working Party was created under the Directive 96/45 of the European Parliament and of the Council of 24 October 1995 (article 30), as an independent European advisory body on data protection and privacy. Article 29Working Party issued Guidelines on Consent under Regulation 2016/679" clarifying requirements of the EU GDPR in April 16, 2018.

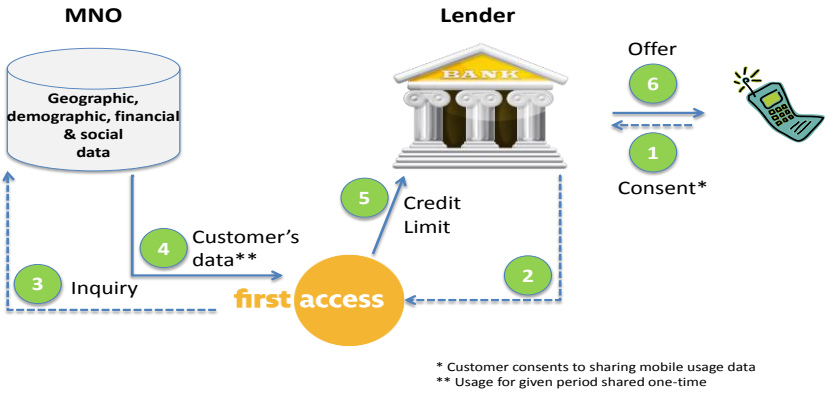
Source; Destacame, BID workshop on Fintech, March 2017

32. **The adoption of consent mechanisms could be done in a cost-efficient manner.** For example, when collecting payment transaction information, the introduction of 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 (CGAP) 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. As an example, "*This is a message from First Access: Mobile phone records are information captured when you use your phone, including phone calls, SMS, airtime top-up, or a mobile money account. Questions? Call First Access 12345678*."[26] However, the research also showed that consumers are prepared to allow use of their data if it means that they can obtain a loan**.**[27] The figure 4 below shows a scoring model using data from different sources where consumers provide consent to sharing mobile usage data.

[26] http://www.cgap.org/blog/simple-messages-help-consumers-understand-big-data

[27] 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/publications/informed-consent-how-do-we-make-it-work-mobile-credit-scoring.

Figure 4- Consent captured through a mobile device (Source First Access, 2017)



III.3. Guidance on Consumers' Rights

POLICY GUIDANCE: Consumer's Rights

Data controllers should enable mechanisms that allow consumers access and the ability to correct their information as well as request the deletion of data -when appropriate- based on applicable laws or rules on retention periods. In addition, consumers should also be able to object to the processing of their information for certain purposes (i.e., marketing). Consumers should also be given the opportunity to transfer their data to any other service provider of their choice without affecting the usability of the data.

Key Considerations

- ***Access to consumers' own data**. At the minimum allowing consumers to access their own data is broadly accepted and practiced in countries where a data protection law is in place and also in those countries that there is no data protection law but there are industries that collect, process and distribute data as part of their core business. Timeline to provide access ranges from 1-7 days.*
- ***Enable the correction of data.*** Consumers should be given options to correct their data. There is typically a timeline between the request from the consumer to the final resolution by the data controller. This timeline ranges from 7-25 days. However, for the use of alternative data from open sources as opposed to closed networks is important to highlight the need to identify the data source and the person responsible for the accuracy of data as such person could also be the responsible party to correct such data and respond the consumer.
- ***Cancel/erase data**. The right to cancel (erase) data is linked to the right to be forgotten, the obsolescence of data and the usefulness of such data. In closed networks (e.g., credit reporting systems), data cancellation is frequently replaced by stopping the distribution of such data or not including such data in the scores to be developed while data is stored in the CRS files for longer periods of time.*
- ***Opposition of data collection** and further usage refers to the choice of the consumer to object the processing of data for certain purposes. This is typically the case of such use for marketing related purposes through the introduction of white lists for example. However, there are certain*

types of data and circumstances where the consumer cannot object the processing of such data (i.e., credit repayment data for credit risk evaluation when such repayment is in default). In closed networks there are certain data items considered mandatory and therefore not subject to consumer choice. In open networks, the choice of consumers regarding the further use of data is broader.

- ***Data Portability-*** *Consumers should be able to obtain and reuse their personal data for their own purposes across different services. They should be allowed to move, copy or transfer personal data easily from one platform to another in a safe and secure way, without affecting its usability.*

Explanatory Text on Consumers' Rights

33. **The big-data effect has often raised concerns about the potential risk to consumer privacy when data from various sources 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.

34. **As highlighted in the recent G20 DFI HPLs, "Consumers also need to have transparent, affordable and convenient access and correction rights**."[28] 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. Consumers may not know who is holding, or has access to, their data, 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.

35. **General Principle 4 of the GPCR 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 of information about them." 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 the concept of portability could provide certain protection to consumers when third parties extensively use their information.

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

36. **The OECD Individual Participation Principle, in summary, provides consumers with the right to find out if a data controller has information about them;**[29] **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.** The Madrid Resolution contains similar principles dealing with the right of access and rights to rectify and delete.[30] 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. 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.[31] 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.

III.4. Guidance on Security of Data

POLICY GUIDANCE: Security

Consistent with existent national cybersecurity plans, industry participants should conduct periodic cybersecurity risk assessments, develop policies and procedures to effectively respond to cyber incidents, communicate cyber incidents to all relevant parties -including consumers- as soon as practicably possible or as required by law, and devote resources to assess, monitor and mitigate consequences of cyber-incidents. These measures are also applicable for any outsourced service that involves the processing, access and storage of personal information.

Key Considerations

- ***Cybersecurity is not a matter of IT only****- Cybersecurity should be part of the overall risk management policies and procedures of any service provider or data provider. In this context, identifying potential threats, enabling mitigating measures and setting up prompt response to incidents would contribute to minimize the consequences of a cyber incident.*
- ***Develop Policies and Procedures to effectively respond promptly to cyber incidents.*** *While a timeline for responding to such incidents should be adopted at domestic level, it is necessary that consumers are informed about such incidents -particularly data breaches- and given mechanisms to mitigate potential consequences for such data breaches (e.g., freeze their credit scores; alert systems when a new product or service is requested using consumers' personal information).*
- ***Legal provisions to ensure accountability for data breaches in case of outsourced service.*** *Including specific provisions in the contractual arrangements between industry participants and third parties service providers regarding security measures and accountability could help mitigate consequences of data breaches.*

[29] 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.

[30] Madrid Resolution, Principle 9.

[31] Convention 108, Article 8

- ***Adoption of Security Measures****. Regulators could encourage the adoption of security measures to avoid data loss, corruption, destruction, unauthorized access or misuse of such data. These measures could also include agreed protocols for incident response including notice of data breaches. While the timeline for such notice varies from one country to another (e.g., EU adopted 72 hours since the incident, while in the timelines vary from state to state).*
- ***Identify a person within the organization to act as Data Security Officer (DSO).***

Explanatory Text on Data Security

37. **According to the G7 (2016), cyber risks are growing more dangerous and diverse, threatening to disrupt our interconnected global financial systems and the institutions that operate and support those systems.** Data is becoming a key asset and personal data and identity theft has become a major risk for consumers. In India, security breaches left 1 billion individuals at risk of identity theft. In the US, 15 million were affected by ID theft in 2016 and in 2017 a major data breach affected 145 million individuals and included information on social security numbers and other personal information which could also put these individuals at risk of ID theft.

38. **Internationally agreed frameworks capture the need for safeguards to protect data against unauthorized access, loss, destruction, and data corruption.** The OECD Security Safeguards Principle establishes: "Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data."[32] The Madrid Resolution places stronger emphasis on security issues, containing separate principles on security measures and the duty of confidentiality. 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.

39. **Adopting cybersecurity strategies faces several challenges**. Awareness and perception of risk differs from one individual to another, which leads to difficulties in identifying cyber threats. In addition, there is reluctance to share information on cyberattacks due to reputational fears. The role of the chief information security officer within financial institutions nowadays takes a more prominent role which needs to be integrated into the corporate governance of the organization. One critical aspect of cybercrime is the non-local aspect when actions can occur in different jurisdictions, simultaneously and activated by remote activity. Finally, this extensive presence of cross-border nature in cybercrimes implies a collaboration between authorities from different jurisdictions and potential legal vacuums.

40. **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 risk; in fact, the potential fallout 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.

[32] Please see OECD Guidelines Principle 11.

Box 4- Lessons learned from Data Breaches

Equifax Case

In 2017, 143 million records of US residents including information on social security number, name, date of birth, address and in some cases drivers' license where accessed by unauthorized persons. This information poses a high threat to Identification theft. Equifax disclosed in July an ongoing investigation of a potential unauthorized access caused by a software vulnerability which happened in March. This data breach resulted in lack of consumers trust, reputational risk to the organization (23 lawsuits have been brought against the company) and potential identification theft.

Reports on the response to the incident indicate:

(i) Slow communication response;
(ii) Lack of adequate governance arrangements to escalate the matter on a prompt manner;
(iii) Inadequate consumer protection mechanisms to verify compromised records;
(iv) Gaps in security measures (risk assessment, vulnerability assessment, patching cycles), and
(v) Outsourced services management deficiencies.

III.5. Discrimination

POLICY GUIDANCE: Discrimination

Policy makers and industry participants should adopt measures to ensure that predictiveness of alternative data is tested and verified, that scoring models developed using alternative data do not unfairly discriminate against protected groups. The use of alternative data that carries forward historical discrimination is either prohibited or restricted, taking into account its ability to predict risk and the availability of alternative decision-making tools.

Explanatory Text on Discrimination

41. **All data used to make credit decisions, whether traditional or alternative, has the potential to result in discrimination against people based on factors such as race, gender, national origin, marital status, etc**. For example, if gender is collected on a loan application, and a lender's scoring system allocates fewer points for women than men, discrimination clearly results.[33] But, it is also possible that a seemingly neutral factor, such as whether a lender considers part-time income, could result in discrimination. Take a lender that is unwilling to consider part-time income in support of an application for a credit card. While neutral on its face, in some economies, women and older people are more likely to work part time. Therefore, some applicants could be disadvantaged based on their gender and age. The fact that such discrimination is a result does not necessarily mean that discrimination was intentional.

[33] Please note that Principle 5 of the HLPDF establishes "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)".

42. **In some cases, traditional credit history disparities are a result of historical discrimination, e.g., such as how certain groups may have been treated in the past which could impact their educational and employment opportunities.** Despite such potential discrimination, some countries are willing to accept some level of discrimination if the assessment of a specific individual is nonetheless highly predictive of their credit risk, especially if no better means of assessment are available.

Box 5- Data analytics discrimination on Minority Groups

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. 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 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). 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". 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.

Source: *New Forms of Data Processing Beyond Credit Reporting*, The World Bank,2018

43. **Alternative data offers much promise including the prospect of bringing individuals into the formal financial system who were previously excluded because they lacked traditional historical data including immigrants, refugees and young people seeking credit for the first time.** However, alternative data also presents the risk of numerous types of discrimination. There may also be unintended consequences from the use of this data. For instance, some government or private sector jobs may involve frequent relocation incorrectly suggesting a lack of stability. The challenge is how to promote use of new data sources while at the same time protecting consumers from unacceptable consequences.

44. **The determination of what constitutes a protected characteristic is a policy matter that will vary from country to country.** Use of various types of alternative data, however, pose significant risks of contravening a country's public policy. Social media provides a good example. If a credit applicant lacks a traditional credit history, a lender might obtain access to that applicant's social media accounts. From them, the lender might conclude that since the applicant's contacts were creditworthy it is more likely that the applicant was creditworthy as well. One problem with this scenario is the reverse – denying credit to an applicant due to his contacts status since those

contacts may share a protected characteristic with the applicant thus perpetuating discrimination against that group rather than considering individuals on their own merits.

45. **In addition to discriminating in approving or denying credit applications, alternative data could also provide a basis for 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. 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). Research conducted in 2015 by the White House and the U.S. Federal Trade Commission has indicated that the use of alternative data may result in discriminatory pricing. 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 higher price while others are offered a lower price available to the rest of the market. Research conducted in 2016 by the Federal Trade Commission also highlighted the concern that "alternative data analytics could affect low-income, underserved populations, and protected groups" especially in relation to credit and employment opportunities.

46. **Transparency is a tool that can help police against discrimination**. If users of alternative information disclose the sources of such data, the data subjects may be able to assess the validity of such information and how it is being used. In addition, a requirement that users of alternative data articulate non-discriminatory grounds for decisions can also serve as a check against discrimination.

47. **The first question regarding use of alternative data is whether it is as predictive of a desired outcome as it purports to be.** If the claim is that it reduces credit risk posed by consumers lacking in a traditional credit history, testing can and should be done to determine whether credit risk assessment using alternative data is at least equal to and ideally better than other available sources of evaluation. If not, then there is far less reason for its use.

48. **The next question is whether use of alternative data results in unacceptable forms of discrimination.** Each country will have the opportunity to decide how to address potential discrimination. One option would be to prohibit the express use of prohibited criteria, e.g., race or gender, in making decisions. Yet, it would still be possible, as noted above, that the use of even facially neutral criteria could result in discriminatory outcomes. This could potentially be determined by retrospective studies or a study of the results of present decision making using such data. Even if proponents of alternative data argue it is hard on its face to assess why certain decisions are made, it would still be possible to examine the results of those decisions.

49. **Finally, if it turns out that alternative data is highly predictive of a desired outcome, yet it also has discriminatory consequences even without the use of prohibited criteria, the question could be posed whether there is another way to make decisions that produce roughly the same results in a less discriminatory manner.** In some cases, this might mean using different

alternative data; in others it might mean concluding that the consequences are too negative to permit a particular use of alternative data. In some cases, where no better alternatives are available, it might mean using data that results in some discrimination but serves the larger interest of financial inclusion. As an illustration of the potential pervasiveness of profiling, the Federal Reserve Board in the United States recently confirmed the evidence of practices where credit card holders had their credit limits reduced and interest rates increased because their spending patterns showed that they had begun to shop at discount stores, even in the absence of delinquency in any account (Rule, 2008; Board of Governors of the Federal Reserve System, 2010).

III.6.- Cross-border data flows

POLICY GUIDANCE: Cross-Border Data Flows

Authorities should enable cross-border data flows when appropriate and consistent with privacy laws and international privacy frameworks. Cross-border data flows should be subject to mechanisms that ensure accountability of data controllers and industry participants. There should be in place procedures and policies to allow consumers to implement their rights - - regardless of where the data is stored or has been transferred. Finally, cooperation agreements between authorities allows for adequate implementation of privacy rules across borders.

Key Considerations

- **Data Localization-** When adopting data localization policies, authorities should take into consideration reduced efficiency from existing infrastructure, and the level of protection provided by other jurisdictions regarding privacy and data. When data localization rules are in place, authorities should still enable the cross-border through consumers' consent or any other legal bases.
- **Conflicting laws-** Authorities should seek ways to develop for a harmonized international approach to common data protection/privacy goals including consumers' rights, dispute-resolution mechanisms, accountability for data errors, and data-security measures.

Explanatory Text on Cross-Border Data Flows

50. **Enabling cross-border data flows facilitates trade and commerce between businesses in different jurisdictions.**[34] However, the World Trade Organization (WTO) General Agreement on Trade in Services allows for restrictions to protect privacy of individuals but also calls for even levels of protection. This is also the approach followed by the EU framework within the Directive 96/45 and then the GDPR that establishes a level of adequacy of protection is similar to the one in the EU. The 2009 Madrid Resolution, for example, encourages consistent international protection of personal data and embraces privacy approaches from all five continents to facilitate "the international flows of personal data needed in a globalized world." But also the industry recognizes that The mobile industry recognizes that data privacy regulation, including rules on cross-border

[34] Article XIV (c) (ii) of the WTO's General Agreement on Trade in Services (GATS) allows trade restrictions that are necessary for "*the protection of the privacy of individuals in relation to the processing and dissemination of personal data and the protection of confidentiality of individual records and accounts*", specifying that "*such measures are not applied in a manner which would constitute a means of arbitrary or unjustifiable discrimination between countries where like conditions prevail, or a disguised restriction on trade in services*."

data flows, is necessary[35]. APEC together with civil society and industry representatives developed in 2011 a set of privacy rules for business to adhere[36].

51. **While there are some countries that have in place comprehensive data protection and privacy laws, others have no rules in place and others rely on internationally agreed-upon frameworks.** The need for standardized common rules to privacy and data protection in economic areas are highlighted by the APEC framework, the OECD Guidelines and finally the EU's General Data Protection Regulation. While each of them present some differences, there are commonalities to all three frameworks. International compatible frameworks will bring certainty and predictability to consumers, businesses and governments.

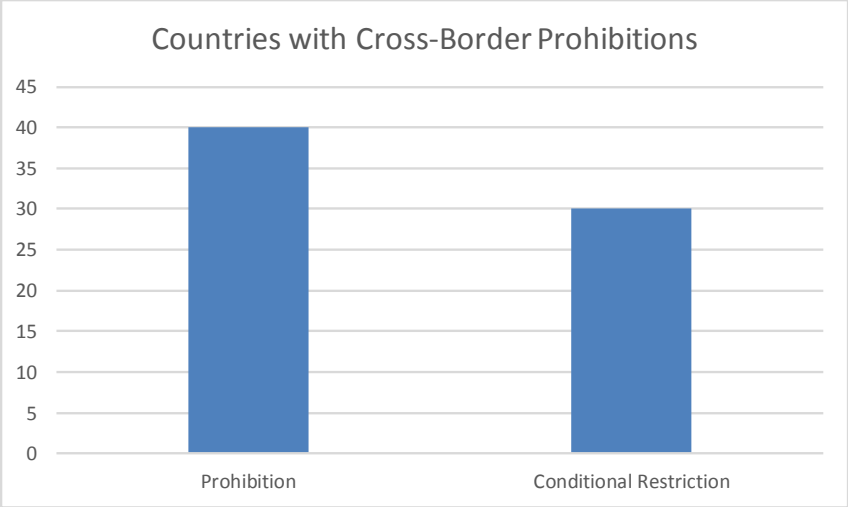
52. **Despite the significant benefits to companies, consumers, and national economies that arise from the ability of organizations to easily share data across borders, dozens of countries—across every stage of development—have erected barriers to cross-border data flows, such as data-residency requirements that confine data within a country's borders, a concept known as "data localization."** Within the EU space there is free flow of data, however the EU has only recognized 12 countries as having an "adequate level of protection" to enable the flow of data from EU citizens to those countries. These countries include: Andorra, Argentina, Canada, Switzerland, the Faeroe Islands, Guernsey, Israel, the Isle of Man, Jersey, New Zealand and Uruguay. The United States developed a framework (safe harbor) to enable the flow of personal data between EU and US although the initial framework failed and was replaced with the Privacy Shield. There are also exceptions to data localization subject to the performance of an existent contract.

53. **Data localization rules require organizations and companies to use local data storage and technology to process data of the citizens of a given country.** There are some countries that include some level of restriction to the storage and further processing of data outside their borders. In this respect, the General Data Protection Regulation (GDPR) includes restrictions to transfer data to countries that do not provide an adequate level of protection (i.e. similar to the one within the EU). This approach is followed by the EU, Switzerland, Israel and Japan.

[35] GSMA "Cross-Border Data Flows", 2017.

[36] The APEC Cross-Border Privacy Rules (CBPR) System, endorsed by APEC Leaders in 2011, is a voluntary, accountability-based system that facilitates privacy-respecting data flows among APEC economies. It has four main components; (i) recognition criteria for organizations to become an APEC CBPR certified Accountability Agent, (ii) questionnaire for businesses that want to be certified, (iii) assessment criteria to be used by APEC CBPR certified Accountability Agents and (iv) regulatory cooperative agreement. There are currently five economies participating to the APEC CBPR system including United States, Mexico, Japan, Canada and the Republic of Korea.

Figure 5- Cross-Border Data Flows Prohibitions



Source: own elaboration based on Data Protection and ICT laws (75 countries, 2018)

Role of authorities

POLICY GUIDANCE: Cooperation and Coordination

Authorities need to develop a National Strategy towards privacy and set up coordination mechanisms between authorities and industry participants at domestic and international level. Ideally, a legal framework on data protection and privacy inspired in internationally agreed-upon framework should be established. Identification of roles (e.g. enforcement, rulemaking and supervision) regarding data protection and privacy of each authority as well as the establishment of formal cooperation mechanisms would allow for smooth implementation. Dialogues should be fostered between financial sector authorities, ICT, consumer protection and data protection and privacy authorities for the appropriate implementation of privacy rules related to the collection, processing and further use of alternative data.

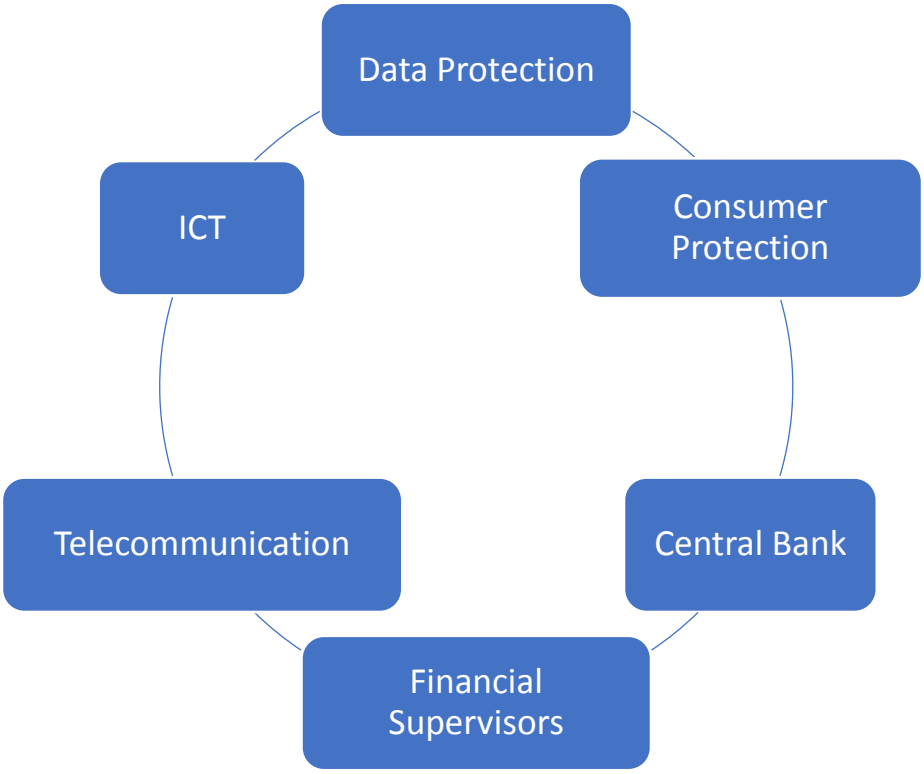
Key considerations

- **Dispute Resolution Mechanism**- Regardless of the types of authorities involved in the enforcement of data protection and privacy rules, it is important to allow consumers access a dispute resolution mechanism that is neutral, agile and affordable.
- **Rules on privacy should be proportionate, consistent at domestic level and certain**. Such rules should take into consideration a functional approach to the activity and foster the application of rules to the same type of activity regardless of the nature of the industry participant.
- **Authorities vested with similar mandates should cooperate towards adequate implementation of privacy rules**.

Explanatory Text for Coordination and Cooperation between authorities and industry participants

54. **In adopting Data Protection measures with the objective of enhancing digital financial inclusion, it might be necessary to coordinate actions between different authorities at domestic level.** The following chart includes the type of authorities typically involved in the adoption of data protection/privacy related measures.

Figure 6- Illustration of potential authorities with mandates on data protection and privacy aspects



References

Cavoukian, Ann “Privacy by Design: The Seven Foundational Principles” (Information and Privacy Commissioner of Ontario, 2011)

European Commission, EU General Data Protection Regulation, 2017

G20 High Level Principles for Digital Financial Inclusion, GPFI, 2016

Grady R. Montes F. and Traversa M., “New forms of Data Processing Beyond Credit Reporting; Privacy and Consumer Protection Considerations”, The World Bank, 2018

GSMA “Cross-Border Data Flows”, 2017.

ICCR, “Policy Brief on Credit Reporting and Financial Inclusion” The World Bank, 2017.

Katy Jacob and Rachel Schneider, “Market Interest in Alternative Data Sources and Credit Scoring,” The Center for Financial Services Innovation, December 2006

Katy Jacob, “Reaching Deeper: Using Alternative Data Sources to Increase the Efficacy of Credit Scoring,” CFSI, March 2006

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

OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data, 2013

Responsible Finance Forum “Opportunities and Risks in Digital Financial Services: Protecting Consumer Data and Privacy”, Berlin 2017.