JUNE 2021

IIF DATA ETHICS CHARTER

INSTITUTE OF INTERNATIONAL FINANCE

INSTITUTE OF INTERNATIONAL FINANCE

Table of Contents

Background and Mission Statement

Data and data-based technologies such as artificial intelligence (AI) and machine learning (ML)[1] are at the core of the transformation the financial services industry has experienced in the last decade. Given its growing importance in the economy, developing ethical principles that go beyond compliance to protect and handle individual customer data[2] is crucial.

The financial services industry has a strong historical record in processing, collecting, sharing, and safeguarding data in an ethical and responsible way – a standing that was evidenced in the Bank of England's 2019 *Future of Finance* report, where a survey cited that 86% of consumers most trusted their financial institution to securely manage their data, in preference to payments providers, tech firms, and online retailers.[3]

However, as traditional sectoral boundaries blur and the financial services ecosystem increasingly features new entrants and tech partners, it is important that all ecosystem participants meet and practice this high standard. This is vital for protecting customers and for ensuring continued confidence in the broader financial system.

Accordingly, the Institute of International Finance (IIF) *Data Ethics Charter* (DEC) outlines a set of principles for the ethical handling of customer[4] data. This Charter articulates the strong practices from across the industry, which all participants in the financial ecosystem, including new entrants and partners should emulate.

The principles-based Charter is intended to be adaptable to fit the evolving policy environment to which IIF membership holds itself. It serves to complement (and not replace) national and sub-national legislation and regulations, as well as international standards.[5]

The principles set out in the Charter do not have binding or legal status, and the IIF does not propose to "police" the principles. However, we encourage all financial sector participants to consider these principles. As the use of customer data and algorithms extends beyond financial services, we invite stakeholders across all industries to incorporate data ethics practices into their activities.

The DEC also helps customers better understand their role through the evolving landscape of data usage in a straightforward manner. Principles and examples of practice provide customers with an overview of how financial institutions responsibly manage, protect, share, and use customer data.

The IIF has developed this DEC with inputs and practical insights from chief data officers and other experts from within our membership across the globe.

[1] Artificial intelligence is the capacity for machines to resemble human intellectual abilities. AI is a broad field with many sub-fields and related fields, including machine learning. ML, a component of AI, provides systems with the ability to automatically learn over time, generally from large quantities of data. For more information, please see the Glossary on page 35.

[2] *Customer data* is defined in broad terms as both the qualitative and quantitative forms of information that is collected and/or created during interactions between a provider of a good or service and the individual customer of a product or service (e.g., a financial institution and its customer). It can include—but is not limited to—personal, behavioral, and demographic data. For more information, please see the Glossary on page 35.

[3] Bank of England, *Future of Finance*, chaired by Huw van Steenis, June 2019

[4] While "customer" can refer to both individuals as well as corporate entities, for the purpose of this Charter, we focus exclusively on the individual.

[5] Some of the principles outlined in the DEC overlap with existing laws in certain jurisdictions.

Introduction

As the financial services sector becomes increasingly digitized and as financial institutions (FIs) increase their use of customer data and data-based technologies such as AI and ML, there is extended focus on data, in its strategic value, and how it is handled and protected.

The ability to collect, store, process, use, and share traditional and non-traditional data is creating opportunities for the financial industry and benefits for customers. Data-driven innovation has led to improvements in credit analytics; customer segmentation, engagement, and insights; client onboarding; automation; product design; fraud prevention; identity verification; product customization and affordability; and the expansion of service to new and underserved market segments.

However, such innovation raises risks related to the ethical use of data, such as unfair bias; lack of transparency; inconsistent application; incorrect assumptions based on analysis of customer data; untraceable source, approach or decision; data quality issues; skill shortages; and data protection, security, and privacy. This could result in reputational harm and loss of trust in addition to the significant resources that need to be reallocated for investigation and remediation of the damages and enhanced risk mitigation in an organization. Thus, the increasing and innovative use of customer data presents both benefits and risks that financial institutions must adequately balance. The increasing criticality of data in the modern economy and the emergence of new data-based technologies brings heightened focus to how institutions can preserve trust, show consideration of human rights, protect customers—including the vulnerable—and ensure that customers' data are handled responsibly from an ethical perspective, while continuously improving their services and offerings to clients.

While more than 120 countries have enacted data protection legislation[6] of some form, there are no global standards to harmonize data protection and governance laws, regulations, and practices in a principle-based, coordinated way. Data frameworks that lack interoperability create gaps in the use of customer data and result in adverse knock-on effects for innovation, poorer quality of products, and higher prices and less efficiencies and options for end customers.

The DEC highlights recommended practices to fit the evolving policy environment. The principles complement and act as a second layer to existing national legislation and regulatory compliance, as well as global standards.[7] Where data frameworks exist, all financial services actors should continue to comply with all applicable laws and regulations. These recommendations offer additional guidance to account for cultural norms and individual company values.

Regulatory and supervisory approaches to data should be based on the overarching principle "same risk, same regulation" and apply to all market players involved, regardless of business model, thus preserving financial stability. Given the interconnection in the digital economy of activities and services traditionally associated to different sectors, ensuring a level playing field generally requires a cross-sectoral approach to data issues.

The DEC focuses on the following areas:

1. Responsible data management cycle;
2. data control;
3. challenges around algorithmic decision-making systems;
4. partnerships and third parties; and
5. skills, awareness, and knowledge sharing.

[6] White Paper - *The appropriate use of customer data in financial services*, prepared by the World Economic Forum in collaboration with Oliver Wyman, September 2018.

[7] To avoid fragmentation in the use of data and the establishment of data usage norms, it is essential to build on the existing hierarchy of data principles: for instance, OECD Principles endorsed by the G20 at a global level, followed by principles outlined by regional and national entities such as the High-Level Expert Group on Artificial Intelligence in the EU, the National Institute of Standards and Technology and the Office of Science and Technology Policy in the U.S., and the Monetary Authority of Singapore's FEAT (Fairness, Ethics, Accountability, and Transparency) Principles.

Complementing each section are examples of practice to help illustrate how the principles presented come into play, and the related experiences and challenges financial institutions encounter as they move to uphold high ethical data standards.

Figure 1: Areas of Focus and Key Principles

Responsible Data Management Cycle	•Manage consent •Increase awareness on data deletion, obfuscation, and/or destruction requirements •Maintain accurate and updated customer data •Ensure data security •Security breach notification and traceability
Data Control	•Data management transparency •Individual control over own data •Manage changes in data use •Customer's data in machine-readable and mobile format •Data mobility
Challenges Around Algorithmic Decision-Making Systems	•Establish appropriate AI/ML governance •Establish accountability •Define fairness •Ensure algorithmic decision-making transparency and explainability •Clearly define human controls •Monitor outcomes •Apply damage prevention measures •Ethics by design verfication
Partnerships and Trusted Third Parties	•Ensure third-party confidentiality •Thorough review of partners accessing customer data •Ensure data privacy through enforceable contracts with partners •Establish accountability through compensation and dispute settlement frameworks
Skills, Awareness, and Knowledge Sharing	•Enhance digital awareness and literacy for customers •Enhance data training and education for employees

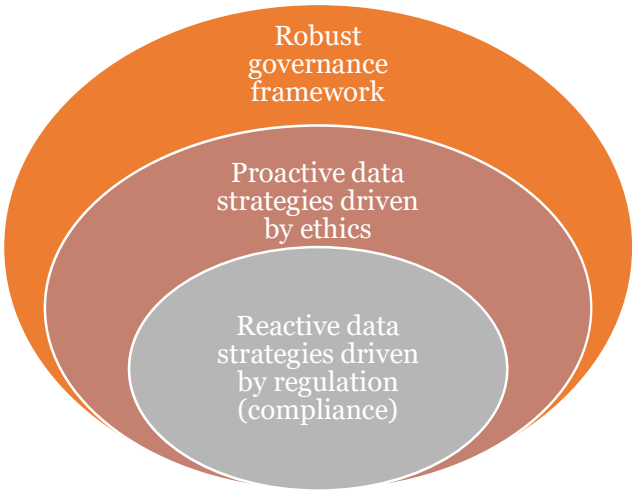
Areas of Focus

1. Responsible Data Management Cycle

The importance of customer's trust and the use of data as an emerging asset brings into focus the ethical use of customer data. Unfair practices are not limited to algorithmic decision-making systems and relate more widely to the use of data – i.e., the bundling, auditing, and compliance aspects of data.

A responsible data management cycle allows for an equivalent of new tools for financial transactions in relation to data – it goes further than compliance to take a more holistic view of data ethics. Figure 2 illustrates the framework as one that builds on the application of laws and regulations that firms have been diligent about (e.g., privacy rules, confidentiality, and data protection laws), includes proactive data strategies driven by ethics, and embeds them in the data governance framework of the firm.

Figure 2: Responsible Data Management Cycle



The inner core is rooted in regulation and represents what firms are already doing (see Appendix A for more details). The middle layer aims at creating proactive data strategies by building a robust internal culture to ask and consider questions that go beyond complying with laws and regulations. And the outer layer combines these two aspects to bring a holistic data management and governance framework.

Proactive data strategies focus on risks and liabilities of using organizational data, to then consider what additional steps should be taken.[8] It addresses the entire lifecycle of data, to include how data is managed, how data is analyzed, and how the insights of data generated in the analysis are used by a financial institution. Data ethics is therefore thought of in every step of the lifecycle of data and informs safeguards and controls that are needed. Ethics-by-design therefore becomes a minimum standard, which specifies the minimum assessment requirements (or questions) that should be considered to give effect to ethics-by-design.

Ethical and responsible management of customer data by financial institutions cannot exist without informing customers in a concise and easily accessible way of how their data is held and used.

The following principles sit in between compliance and proactive data strategies, and are the cornerstone of robust data ethics, and need to be reconsidered under the new evolved framework. In many cases they are integrated in existing data governance frameworks, although they may vary across jurisdictions, depending on the privacy culture, customer expectations, and the regulatory environment.

[8] Proactive data strategies can also stem from data privacy laws, such as privacy-by-design and privacy-by-default. Certain privacy laws, such as the GDPR, require a privacy-by-design control to be implemented. This control identifies and mitigates privacy risks in the design stages for any new processing activities undertaken using personal information (e.g., changes to an existing internal business process, which may include a new or enhanced IT application/system; data analytics or modeling; or the creation of a new product/service).

Manage consent[9]

Financial institutions should develop and regularly adjust technological solutions to allow customers to better manage consent and their data. Customers should be able to grant or deny consent to financial institutions gathering, using, sharing, or storing their data. This includes requesting that their data is no longer used, shared, and stored. Exceptions, including, legal and regulatory requirements, apply.

Generally, financial institutions consider any non-public information to be highly sensitive and worthy of protection. The barriers to overcome before any such information is disclosed without the customer's consent are high, and often limited to cases of law enforcement, prevention of financial crime, legitimate interest, public interest, and regulatory reporting and oversight.

Consent is one of the possible legal bases of legitimate data usage, but not the only one. Outside of existing laws and regulations, consent is related to data ethics as it relates to privacy, ownership, traceability, and misuse of data. In some cases, explicit and traditional consent is not always possible; proactive strategies should develop and regularly adjust technological solutions to allow them to better manage consent.

Increase awareness on data deletion, obfuscation, and/or destruction requirements

Financial institutions should appropriately inform customers of the applicable rules and modalities of destructing data related to them, and customers should be able to access mechanisms to delete their data, if legally applicable.

Customers should also be able to have their data erased in accordance with applicable law. In practice, many customers are not aware that they can revoke their consent and may not know how to access the mechanisms to restrict how their data is being managed.

Principles at Play: *Consent and ethics in conflict - offering savings account for customers with newborns based on transaction data*

If transaction data shows that a customer starts buying diapers, a financial institution can detect a change in the private life of the customer and identify the customer as a new parent. Can the FI proactively offer a life insurance or a savings account for the newborn? And if so, would the new parent appreciate this, or rather consider it unnecessarily intrusive?

The assessment of this situation from an ethical point of view can significantly differ in various countries, often due to their specific circumstances and variations in culturally acceptable behavior. In some countries, sending customers a personal offer based on transactional data may not be considered problematic; but it may be considered unethical even with the customer's consent for monitoring.

From another perspective, leveraging transaction data to help customers can be a way for FIs to empower their clients if consent has been provided by the customer and the communication reflects the purpose and the origin of the data used.

To prevent being perceived as intrusive, FIs should assess carefully and according to country-specific practices and customs what is ethical/unethical and develop processes for aligning use cases to ethical principles.

[9] Consent is defined as any freely given, specific, informed, and unambiguous indication of the data subject, i.e., the customer's wishes by which the customer, by a statement or by a clear affirmative action, signifies agreement to the processing of data relating to him or her. Consent is important because it shows that a customer understands what is being done with their data.

Maintain accurate and updated customer data

Financial institutions should strive to have updated and accurate data on their customers, by obtaining data from trusted sources, regularly checking their correctness, and deleting it when the purpose for their processing has expired.

Having accurate and up-to-date information on customers requires regular checks and the maintenance of coherent datasets. This includes maintaining good quality data on customer transactions/behaviors and providing a mechanism for the customer to correct their personal information/data. This helps financial institutions meet KYC (know your customer) / due diligence requirements, better understand and protect their customers and their financial dealings, and manage risks more prudently.

Principles at Play: *Zurich's Data Commitment*[10]

Zurich Insurance Group has introduced a Data Commitment across their entire group that promises to honor customer's long-standing trust in sharing their data with them. Zurich has committed to voluntarily achieving high ethical standards, going further than what is required by rules like the EU's General Data Protection Regulation (GDPR) on usage of AI, Big Data, and other forms of advanced analytics.

Zurich made four promises as part of their Data Commitment pledge:

1) to keep their customers' data safe;

2) to never sell customer's personal data;

3) to not share personal data without being transparent about it; and

4) to put their data to work so Zurich can fuel new ways to improve their customer's lives.

Zurich has established the Cyber Fusion Center to protect their customer's data by combining cyber threat intelligence, response, forensics, and vulnerability management teams.

Ensure data security

Financial institutions should develop and implement appropriate safeguards to ensure the security of the customer data they hold. Responsibility, however, should be balanced and shared between FIs and customers.

While data security has always been a primary concern for FIs, it is becoming increasingly important as the financial sector is one of the largest targets of sophisticated cyber-attacks.

FIs should be responsible and accountable for implementing appropriate data protection and security frameworks and good practices to prevent events such as breaches, misuse, etc.

Data security responsibility should be balanced and shared between companies and customers, with the latter group also playing a key part in facilitating responsible and secure data procedures. Enhancing digital awareness and literacy for customers can help customers be more careful with their data and accept software updates or changes offered through financial institution channels.

[10] https://www.zurich.com/media/magazine/2019/earning-trust-to-unlock-the-power-of-data

Security breach notification and traceability

Customers should be notified in the event of a data breach having sensitive information on them and be provided with information related to improper use or access. Also, customers experiencing or suspecting the occurrence of such cases are encouraged to report them to financial institutions. Breach notifications depend on the domestic legislation applicable.

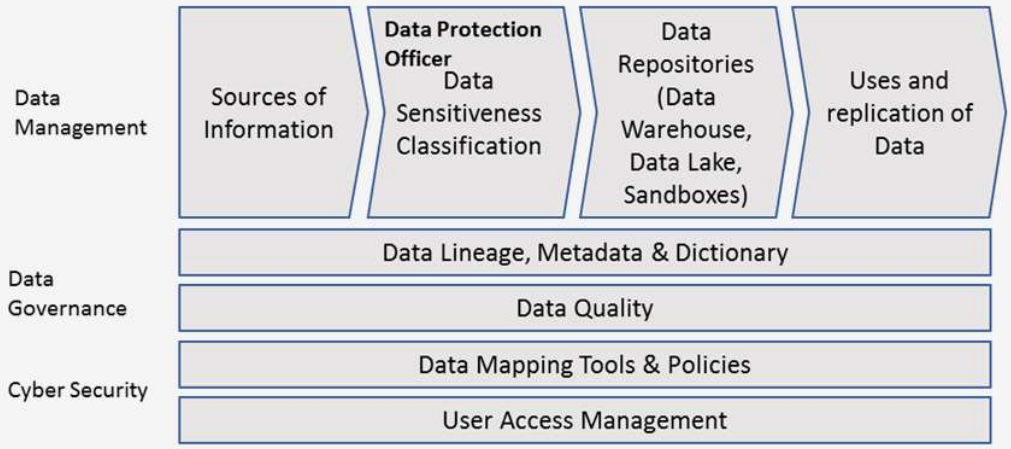
***Principles at Play**: Banco de Crédito del Perú (BCP)'s data protection framework to cope with digital challenges*

With more than 130 years of business, BCP considers data as one of its main assets and devotes significant effort to continuously refine data protection practices to minimize losses or security breaches.

Prior to the digital era, data protection usually relied on operational measures to avoid exposing printed data by employees and partners/vendors. The current digital environment, however, presents a more complex situation regarding management of data, with the exponential increase and use of available data, thus BCP developed and implemented a new data security and protection framework.

BCP's framework (Figure A) was developed firstly by mapping all data sources, grouping them by domain and cataloging the information available to the organization. Mapped information was then classified based on sensitiveness, i.e., identifying PII (Personal Identifiable Information), and high sensitivity data. Once mapped and grouped according to the level of sensitivity, BCP identified the location of critical information – mapping not only to the central and formal data repositories of the Bank, but also all the informational, individual businesses' sandboxes. Cyber security tools were used to track entire servers and computers to find users, access points, and repositories with sensitive information. Under this framework, the creation of the position of the Data Protection Officer was fundamental to guarantee that all applicable legal requirements and data protection policies are respected by the organization and its

Figure 3:BCP's Data Management Framework



The first results revealed that 60% of the thousands of employees that had access to sensitive data, did not need to access that data. Additionally, those same individuals had been creating other data repositories, with no data access control to those data. With the implementation of this framework, access to sensitive data, and the issue of data replication without control was reduced to less than 30% of previous figures. Additionally, the framework was implemented without affecting the continuity and provision of business operations.

Lastly, BCP created a fresh new data dictionary, supporting the entire community of users of information, as well as the lineage of information. As a result, BCP is now able to map the source of each datum and identify all related customer information available in the organization. This way customer data can be handled

coherently and more flexibly. For example, in case a customer requests not to use his/her data, this can be implemented throughout the data management system in a reliable way.

2. Data Control

Customers of financial products or services should understand how data related to them is used, shared, and moved around.[11] Unless customers know how their information is being used, they will be unable to determine what control they wish to exercise.

Data management transparency

Financial institutions should provide customers of a financial service or product with clear, concise, and accessible information about why their data is collected and how it is stored, used, safeguarded, and shared. This is commonly achieved at the time of data collection via privacy statements or notices unless data is collected via third parties in which case notice requirements can vary by jurisdiction.

To ensure transparency and user-friendliness, it is important that FIs carefully define standards for how and why the data of customers is collected, stored, used, safeguarded, and shared with third parties not acting on behalf of the financial institution.

Additionally, firms should strive to explain their processes in a clear, transparent, and simple way, through accessibility and usability techniques that can be understood by customers that are not technology experts.

Individual control over own data

Customers of a financial product or service may have the right to control how data related to them is used and/or shared with third parties not acting on behalf of the financial institution. The extent of these rights and the ability of customers to exercise control will vary, depending on the type of data (e.g., obligatory, optional) and the processing purpose in question. Many firms have already committed to providing customers with control over the use of their personal data wherever possible and operationally feasible; for example, by providing opt-outs (or refusing to opt in) for data being used for marketing, data analytics, and web analytics.[12]

Certain types of client data are required by the financial institution as part of a specific product/service offered. Disclosures should provide transparency around what products or services a customer may miss out on should they choose to opt out.

Firms have also redesigned user interfaces with privacy by design, i.e., the user experience (UX) and user interface (UI), by building interactions that are clear, transparent, user-friendly, and easy to understand. Interfaces are being designed to enhance data settings making privacy settings more accessible and simpler both in design and language. Customers rarely find websites today without a privacy statement and information panel on the use of cookies that explains how and why that particular company collects, uses, and shares a customer's data.

[11] There are limits to this control, for example financial institutions can process and share customer data for AML purposes. These rights will also vary across jurisdictions.
[12] In certain cases, it may occur that customers can opt out, but then they cannot get the product/service offered (e.g., without the customer's data, the bank cannot properly assess a customer's probability of default).

Principles at Play: *Control over own data – offering new car loans to customers based on car registration number*

A customer pays for car insurance via a bank where the car registration number is part of the transaction data. Based on the registration number, the bank calculates the age of the car, and determines that the customer could be interested in a loan. Should the FI offer the customer a new car loan if the processing of data was collected for a different purpose?

In this case, the data collected includes both transactional data and personal data (i.e., the car registration number), and the re-use of personal data could be perceived as intrusive.
The FI needs to consider whether applicable regulation – in this case the GDPR - allows the re-use of processed personal data for another purpose. Under the GDPR, personal data can only be re-used for a purpose which is compatible with the original purpose for which the data was collected, therefore in this case, it is likely that additional consent from the customer would be required for re-use of this data, giving the customer control over their own data.

In practice, FIs will usually collect customers' marketing preferences on account opening and this agreement between the customer and the FI may enable such marketing to proceed.

Manage changes in data use

Financial institutions need to update customers on modifications in data use and sharing and their implications in a timely manner, and communicate those changes in an easily understandable way.

Purposes of data use and sharing may change over time, which means that financial institutions should also take care that such modifications and their implications on customers are communicated in an appropriate and timely manner. As detailed above, FIs must always ensure compliance with any applicable regulation, such as the EU's General Data Protection Regulation (GDPR), before modifying data use and this may require customer consent.

Customer's data in machine-, human-readable and mobile format

Financial institutions should strive to make sure that the portable aspect of customers' data is in a structured, machine-readable and human-readable format that facilitates its mobility between data platforms and controllers. Exceptions, including, legal and regulatory requirements, apply.

While there are certainly considerable operational challenges involved, financial institutions should strive to make sure that their customers' data is in a structured[13], machine-readable and transferable format as this facilitates its mobility between data platforms and controllers and empowers customers with greater control. In some cases, this is a regulatory obligation, for example, in relation to personal data processed under GDPR.

Data mobility

Customers should have control over their data (across all sectors) and be able to share it with third parties in an automated way (considering the options allowed by applicable data sharing regulatory frameworks).

[13] Structured data refers to data that is well organized and clarifies and standardizes the relationship in the data.

The principle of data mobility seeks to ensure that an individual's data is not siloed or walled off by gatekeepers and can be easily moved around. Data mobility is becoming an increasingly important concept with more and more jurisdictions around the world—including the EU, Canada, Australia, Brazil, Singapore, Thailand, the UK, and California—starting to recognize individuals' right to control their data and have it easily transferred from one commercial entity to another at the request of the customer (see Annex A for more information).

Principles at Play: *Barclays commitment to simple data mobility for its customers*

Barclays, one of the largest banks in the UK, is committed to ensuring easy data mobility for its clients.

Customers simply use the bank's online *Right to Data Portability* form to ask the bank to send an electronic summary of the personal data they have shared with the bank previously. Upon completion, the bank may contact the customer for security purposes. Once the request is approved, it will typically take up to a month for the bank to process it, though for more complex requests, it may take up to three months. The data is sent to the customer online, in a secure way, and in a machine-readable format. This means the data will be provided to the customer in a spreadsheet that can be opened and processed by a computer so that the data can be easily read by other companies or organizations the customer wishes to share it with.

3. Challenges Around Algorithmic Decision-Making Systems

Clear ethical principles on the way data and technology are used are critical to the Charter.

The volume of data available combined with the new technologies, such as artificial intelligence (AI) and machine learning (ML), can deliver customized products and services to customers in a faster and more agile manner, reduce processing times, and ultimately increase customer satisfaction. For industry players, these new insights can make them more efficient and enhance the products and services they provide.

Within the financial services industry, algorithmic or automated decision-making systems[14] are used for a variety of applications from credit decisions to marketing, financial crime detection and surveillance, customer assistance and advice, and insurance eligibility. Outside of financial services, we see these at play in streaming service recommendations (e.g., Netflix, Hulu, etc.), shopping recommendations (e.g., Amazon), and within the public sector (e.g., criminal justice sentencing and probation decisions). The implications of using algorithms to guide or automate decisions very much depend on their use case, as these can differ among sectors (from streaming service recommendations to healthcare diagnoses) and intra-sector (from marketing to credit decisions).

While benefits are gained, there are also challenges and risks that need to be addressed around unfair bias and unfair outcomes, [15] and around the usage of organizational data.

Challenges related to using organizational data include *data accessibility* (i.e., who has access to the decision-making) and *comprehensibility* (i.e., who understands the relevant aspects of the larger modeling process), which overlaps with being able to trace decisions.

Establish appropriate AI/ML governance

Financial institutions should establish enterprise-wide internal governance frameworks across business functions with clear principles of accountability to provide transparency around the ethical

[14] Please see Glossary for a definition of algorithmic decision-making systems.
[15] For more information, please see two of our previous papers: IIF, *Machine Learning Thematic Series: Bias and Ethical Implications in Machine Learning*, May 2019, https://www.iif.com/Publications/ID/3338/IIF-Paper-on-Bias-and-Ethical-Implications-in-Machine-Learning; and IIF, *Machine Learning Thematic Series: Machine Learning Recommendations to Policymakers*, September 2019, https://www.iif.com/Publications/ID/3574/Machine-Learning-Recommendations-for-Policymakers.

use of data. This includes developing processes and risk management protocols to test, monitor, and govern data to minimize biases, inaccuracies, and unintended consequences.

A proper model governance framework must be designed to ensure that algorithmic decision-making systems follow robust measures for model development, implementation, and use, as well as sound processes for model validation.

This first includes measuring the accuracy and relevancy of the data (i.e., does the past data represent the future accurately?), and the accuracy of the model fit (i.e., does the model distill the relationship in the data accurately?). Data relevancy and representativeness should be assessed to determine if it is representative of the population to which the algorithm will be applied. While numerous data strategies exist to manage and protect data, the use of data in algorithmic decision-making systems brings additional challenges, such as inherent data bias.

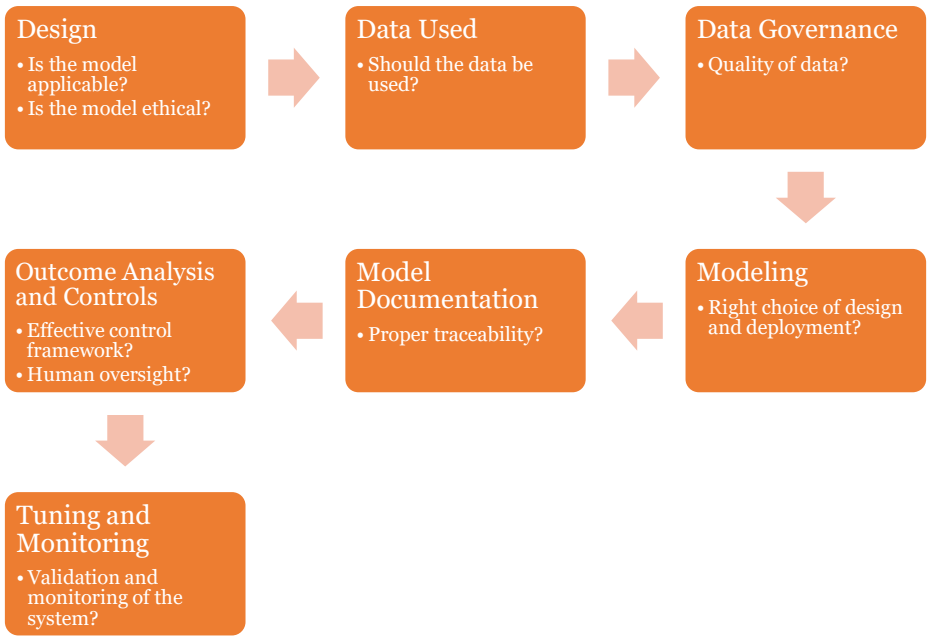
Secondly, FIs should consider fairness of such algorithmic decision-making systems as defined by each institution (i.e., are the encoded relationships desirable?). Bias mitigation measures should be considered throughout the different stages of the model governance process by establishing various mechanisms and controls.

Figure 4 presents a simplified example of the different stages of the model governance process in which unfair bias could be checked and are tied to the following principles.[16] A prior assessment of the usability of the data should be undertaken to ensure that it meets the legal and regulatory, data quality and security requirements.

Regulatory initiatives should take a risk-based approach with appropriate controls that are commensurate with the risk of each specific use case. Regulatory initiatives should be aimed at ensuring high risk applications have a positive impact and reducing obstacles to experiment with or apply new ML techniques. It would be FIs' responsibility to calibrate their actions and requirements under their own internal voluntary codes of conduct for AI with specific transparency requirements.

Regulatory initiatives should remain dynamic, technology-neutral, and futureproof. Regulators should not impose overly prescriptive requirements, rather they should offer a level of strategic tolerance.

Figure 4: Potential areas in which unfair bias can be introduced in the model governance process



[16] IIF, *Machine Learning Thematic Series: Bias and Ethical Implications in Machine Learning*, May 2019

Principles at Play: *DBS' PURE Framework*[17]

DBS Bank in Singapore developed and started implementation of their PURE (Purpose, Unsurprising, Respectful, and Explainable) framework in 2018 as part of a bank-wide effort to elevate data management and governance standards. PURE represents an intuitive approach to assess whether an intended data use case by DBS is responsible or risks creating mistrust. To date, DBS has now trained up to 15,000 employees on PURE, representing around 50% of the company's total workforce.

PURE is the second control gate of DBS' Responsible Data Use Approach. The approach starts by considering questions of data use in a continuum to ensure that DBS gets the right outcomes:

1. Can I use the data (i.e., meet legal and regulatory, data quality and security requirements)?

2. Should I use the data (i.e., PURE)?

3. How I use the data (i.e., model governance)?

The baseline of every use case is the first gateway of data hygiene—is the relevant data good quality, secured, and does it conform to regulatory and legal aspects. Then, those presenting the use case must articulate whether the use case is PURE - purposeful, unsurprising, respectful for both customers and employees, and explainable. Finally, the use case goes through the model governance process, which includes development, testing, validation, implementation, and ongoing monitoring of the model.

PURE involves senior leaders from each business unit with an established operational process, the self-assessment approach under PURE has provided business units with the autonomy to use data responsibly to develop and quickly bring to market innovative products and services without compromising consumer trust.

The review process allows DBS to understand the purpose of data use, avoid any unsurprising effects on customers, respect social norms of the societies in which they operate, and ensure that outcomes are explainable.

When a non-PURE use case is identified but the use case is deemed worthwhile to pursue, it is subject to a cross-sectoral Responsible Data Use Committee review involving senior management, and considers amongst others, individuals, corporate, sustainability and local country considerations. The Committee ensures that non-PURE use cases do not get approval and/or that they are tweaked or enhanced to ensure sufficient mitigating measures in place to meet the PURE principles. Once a use case is approved, it is white-listed, allowing similar use cases to get approved in the future. However, whenever new data is injected, it necessitates bringing back the use case to the Committee. The approval process is use case specific; the Committee meets periodically but for urgent requests they hold ad hoc meetings.

Having a respectful view of the use of data requires DBS to think of diverse societal perspectives on whether you should use the data or not. In this sense, having a broad section of perspectives in the Committee has enabled DBS to also have a broad set of perspectives; for example, having a country perspective is crucial as what may be respectful in Singapore may not be in Indonesia.

DBS has also made a leadership commitment to address the need for increased and continued attention on data ethic at board-level. DBS has embraced accountability for artificial intelligence (AI) and machine learning (ML) models by including an AI expert on their board of directors to "add an additional check and balance."

Although PURE may not work for all financial institutions, it has allowed DBS to gain a competitive advantage in developing a framework that can be applied widely and evolves well in time. As a risk management framework, PURE has clarified and helped to expand the safe space which DBS can use data.

The following principles aim to be a foundation for achieving an ethical use of data.

[17] Contributors: Sameer Gupta, Managing Director, Chief Analytics Officer, DBS Bank; Jeffery Lee, Managing Director, Legal and Compliance, DBS Bank; and Chee Kin Lam, DBS Managing Director and head of the bank's Legal, Compliance, and Secretariat group

Establish accountability

A clear accountability framework should be in place to safeguard the safety of algorithmic decision-making systems. This framework should be based on model governance tools to enable traceability and interpretation of decisions taken in the design process, of data used to train a model, and on all elements needed to make subsequent analysis and inquire if the model deviates from its initial purpose.

Financial institutions need to establish a clear chain of command on data ethics related to algorithmic decision-making systems, this means that the framework should be based on model governance tools and processes to test, monitor, and govern the safety of algorithmic decision-making systems around the ethical use of data. This also includes clearly identifying the ownership roles internally, i.e., the people who are accountable.

Traceability focuses on maintaining records of data characteristics used to train the model, such as data sources and data cleaning, record of the code, and the decisions taken towards deployment, which can help analysis into the outcomes of an algorithmic decision-making system.

For optimal results other factors need to work together in establishing accountability around data ethics in algorithmic or automated decision-making systems, these are: *the volume of data* (the amount of data used), *the quality of data, the diversity and representativeness of system engineers and data scientists,* and *the outcome of the algorithmic decision-making system*. These factors are described in more detail in Appendix A.

To reduce the impact of potential biased outcomes, legislation often limits access to data[18], deterring assessments to avoid discrimination. Allowing access to more data would allow far more representative datasets and would help address unbalanced datasets, and likely lead to more accurate predictions for subgroups. Thus, policymakers should be aware of the interaction of certain rules on innovative technologies.

The use of algorithmic decision-making systems can involve partnerships with third parties, FIs should test their current practices to ensure transparency around model interpretability. Depending on the use case, FIs may include contractual clauses for third parties regarding testing methodology, explainability of results, and/or intellectual property rights which may derived from use of the system.[19]

Define fairness

Financial institutions should adopt a definition of fairness to guide its implementation, which applies to both the inputs (data sets), the process, and the output of the model.

It is crucial for financial institutions to have a definition of fairness, not only at the level of each FI, but also often at the level of individual use cases or groups of use cases. In many cases the definition used may need to vary by country or even by state due to legal reasons. The definition of fairness will then guide their implementation and monitoring.

Algorithmic decision-making systems use historical data to make predictions about the future, thus the accuracy of the historical data and any other input data should be understood as it influences prediction and bias. Biases refer to the overrepresentation of certain data over others. Underrepresentation often leads to poorer performance of certain segments. Though the system may provide benefit to a certain segment, it may disadvantage others. Certain biases are rooted in trends or behaviors that are true, while others may be over

[18] More specifically, GDPR art.9 prohibits the recording of sensitive personal features, posing a barrier when FIs need to use them to ensure that discriminative outcomes are being avoided. See https://gdpr-info.eu/art-9-gdpr/
[19] Wharton University of Pennsylvania, *Artificial Intelligence Risk & Governance*

exaggerated. To combat this, human operators must be carefully trained to maintain the balance between accuracy and effectiveness in the model without over influencing the decision.

A focus should be placed on monitoring outcomes to ensure that they do not lead to unfair discrimination and unfairly biased decisions or unintended consequences. Financial institutions should take measures to govern data and algorithms and monitor the practical implementation of internal policies on safeguarding data and employees' adherence to them. We discuss this in more detail on Appendix A.

Ensure algorithmic decision-making transparency and explainability

Financial institutions should indicate in an easily accessible, plain, and understandable language and readable format when a customer is engaging with an algorithmic decision-making system versus a human. There needs to be transparency towards customers in terms of the purpose, data use, and the outcomes of algorithmic decision-making systems. If the impact of the use case requires it, decisions should be made explainable in a way that can be easily understood by humans. [20]

Algorithmic decision-making systems rely on both traditional statistical techniques and new(er) ML techniques to improve accuracy and fairness. Newer ML techniques are often more accurate than traditional techniques but can be less interpretable. However, some older algorithms can appear more interpretable, but the subtleties can be hidden in the correlations in the model and data. Newer algorithms do not offer those same, misleading shortcuts that require more brute force methods for interpretation.

Transparency towards customers in terms of explanations of the decision outcomes can be achieved in both cases, in the latter it requires additional development efforts to provide explanations that can be easily understood by a human.

Explanations provided to different stakeholders—such as business partners, regulators, or customers—can have different levels of depth, but in all cases, they should be easy to understand and should present the relevant aspects without overly complex explanations. This can be achieved through techniques such as global surrogate models, local explanations, or contrastive explanations, etc.[21] At the same time, achieving transparency should allow financial institutions to protect their industrial know-how and IPs.

It is crucial to test specific parameters carefully and continuously on outcomes to ensure that the results remain valid and can be used as expected. To meet this objective an impact assessment should measure the impact on business continuity, and on customer's basic human rights, including dignity, equality, and freedom. The latter means identifying negative impacts on customers and evaluating risks of unfair bias or other unintended outcomes.

The degree of transparency should be commensurate with the impact of the use case, aiming at ensuring that highly impactful use cases are given a higher degree of explanations and transparency. Additionally, in certain cases transparency cannot be provided to a customer, such as in cases related to fraud or financial crime. Challenges around the usage of organizational data can be addressed by building transparency and accountability into the system. Establishing accountability can be done through an accountability framework that is based on model governance tools that can help trace decisions taken in the design process and data used to train a model. Delivering transparency of outcomes must be done with the customer in mind, that is if the logic of an outcome is too complex for a customer to understand, then helpful explanations need to be provided in context, while avoiding information overload.[22]

[20] We define the concepts of "explainability" and "interpretability" in the Glossary.
[21] For detailed information on this topic, see IIF Thematic Series Paper: *Explainability in Predictive Modeling,* November 2018: https://www.iif.com/Publications/ID/1423/Machine-Learning-Paper-on-Explainability-in-Predictive-Modeling
[22] This touches on the concept of comprehensibility, and links back to the concepts of interpretability and explainability of ML models, which we define in more detail in the Glossary. Comprehensibility or who understands the larger aspects of the model development process varies depending on the different stakeholders.

In this context the difference between inherently interpretable models and post-hoc explainability techniques is important. Inherently interpretable models are designed to be explainable, while post-hoc explainability tools apply interpretability methods after the training, i.e., tools are used to extract information from learned (more complex) models and attempt to highlight the salient features of the model. Approaches to achieve explainability are still evolving, and each technique has limitations and caveats that users need to be cautious about.

Principles at Play: *Standard Chartered partnering with Truera to tackle unjust bias in AI-assisted decision making*

Artificial Intelligence/ Machine Learning ("AI/ML") is increasingly being used to help financial institution (FIs) design better products, improve turnaround times, detect fraud, ensure more efficient and accurate risk management, and deepen financial inclusion.

However, there are potential downsides too; AI/ML models can be inherently opaque and create unintended bias in decision making. It is important to ensure that the sources and causes of such biases are identified, and incidence of biases in the outcomes of using AI/ML are minimized.

Standard Chartered Bank is an active proponent of the use of AI/ML and advanced analytics to improve productivity and better support clients and stakeholders, while ensuring fairness, ethics, transparency, and accountability. To that end, the bank recently partnered with U.S.-based start-up Truera to pilot a solution to improve model quality and increase trust by analyzing models and helping to identify and eliminate unjust biases in the decision-making process.

The bank tested the solution on one of its credit decisioning algorithms which uses a combination of traditional data, and with the clients' consent, alternative data. The solution enabled the bank to analyze the behavior of this model, measure the existence of any bias, assess whether such bias could be justified, identify the drivers of any such unjust bias and decide on any corrective action that might be required.

Standard Chartered is sharing this experience with other industry players through the Monetary Authority of Singapore's Veritas consortium. The bank hopes that this will contribute positively to the discussion around responsible use of AI/ ML in financial services.

Principles at Play: *The use of wearables data in life and health insurance*[23]

The use of wearables data sheds light on the importance of appropriate governance of AI/ML. Several insurers have implemented the use of wearables data to identify customer needs, define new customer segments and develop new policies to address specific needs.

In underwriting, wearables provide data that allows a better assessment of health risks and the implementation of pricing models that incentivize healthy lifestyles. Wearables data can be used in claims to initiate and speed up the claims process, reduce the need for burdensome documentation by policyholders, and identify fraudulent claims.

Wearables data allows the implementation of proactive risk management and early intervention mechanisms and is likely to enhance the accuracy of risk assessments. Customers typically benefit from premium rebates and reward programs for sharing their wearables data with the insurer. However, it may raise concerns of discrimination as health-related lifestyles often correlate with income and education level. Additionally, certain non-demographical variables could create bias in the accuracy of the risk assessment coupled with the data gathered by the variable. Individuals with high health-related risks may face high and potentially unaffordable premiums, that may limit their access to basic medical provisions, leading to a further deterioration of their condition.

[23] Big data and insurance: implications for innovation, competition and privacy, The Geneva Association, page 33, https://www.genevaassociation.org/sites/default/files/research-topics-document-type/pdf_public/big_data_and_insurance_-_implications_for_innovation_competition_and_privacy.pdf.

Therefore, the management of bias is critical to consider how data is processed, as well as how employees are trained to look at outliers to prevent bias and improve automated processing.

Clearly define human controls

The degree of autonomy of an algorithmic decision-making system should always be clearly defined, and the different levels of human control over these systems must be clearly identified based on the specificities of each use case.

Optimal decisions suggested by an algorithmic decision-making solution do not always have to be applied. Its application depends on the level of automation of the system, and the final decision should be based on validation and continual testing of outcomes. In fact, one benefit is that new information uncovered by the algorithm can be adjusted to reduce errors, given the algorithm's ability to learn over time.

Using human interaction in the validation and calibration of ML models can be a double-edged sword. For example, bias can come from humans deliberately inputting abusive language to a chatbot designed to learn dynamically, resulting in it learning to say offensive things. Humans can validate the results of the model to determine the accuracy of its inputs, however this could introduce a level of human bias into the system. In contrast, isolating the ML models would result in inaccurate outputs since the model will only adjust to the trends identified in the data.

Thus, it is important to strike a balance in the level of human intervention involved to limit biases in decision making. A combination of human and machine intelligence to create a continuous loop for training, testing, finetuning, validating and monitoring ML algorithms is increasingly being used as a safeguard. We discuss Human-in-the-loop (HITL) in more detail in Appendix A.

Principles at Play: *Developing tools to enhance explainability of algorithmic recommendations*

A large European bank uses ML models to predict which clients of its retail business have more probability to contract an investment fund. Those models are classification tools that are based on logistic regression, random forest, and gradient boosting trees amongst others.

The challenge in this case is to achieve accurate explanations on which of the 389 variables that conforms the input of the models are more important for its outcomes in a general case and for a sub-segment of customers (i.e., senior vs. junior clients). To solve this, the bank has used, adapted, and extended an open-source library called ALIBI, that includes various methods to explain single observations and to provide feature importance. This tool has been integrated in the analytical platform of the bank, and in the near future will be used for the generation of investment recommendations and incorporated into the product worksheet. This will allow the personal advisors to provide their clients sound reasons on why a given investment product is suggested according to the customer's profile.

Monitor outcomes

Once implemented, financial institutions should control and monitor algorithms to identify, verify, and correct any potential biased outcomes in the ML model operation, especially after every recalibration or training iteration.

Apply damage prevention measures

Creation of risk assessment and management procedures aimed at applying measures that prevent harm or bias from occurring, identifying critical situations, and prototyping risk scenarios using algorithmic decision-making systems.

Ethics by design verification

During design, financial institutions should create mechanisms to implement a set of procedures that identify, verify, and correct any "wrong" or potentially biased decisions. For high impact use cases where more advanced and less interpretable algorithms are used, additional efforts on explainability would be needed.

Principles at Play: *Automation of claims handling processes in insurance – Fukoku Life*[24]

In February 2017, Japanese life insurer Fukoku Mutual Life introduced an AI application to boost its operations efficacy in medical claims processing. The application was tasked with calculating accurate payouts based on details of the administered procedure, period of hospitalization, medical history, and insurance conditions. The application accessed medical certificates, hospital bills, and internal claims history files and scanned the insurance contract for special coverage clauses to prevent payment oversights. It calculated the pay-out and submitted it to a member of staff for review and approval.

The application increased productivity by 30% and yielded improvements in accuracy of pay-outs. Additionally, customer satisfaction scores increased through reduced lead time of pay-outs.

4. Partnerships and Trusted Third Parties

Any market player in the financial industry - from traditional banks and insurers to newer innovators in the open banking space - should uphold strong ethical groundings of fairness and model outcomes when sharing customer data with third parties.

The rise of financial technology firms (commonly referred to as "FinTechs") and digital banks, and the demand of customers for faster and more intuitive products and services, have shifted the ways traditional financial institutions (e.g., banks and insurers) do business. Financial institutions have taken more proactive data strategies that have pushed them to collaborate with trusted third parties to enhance customer experience allowing them to offer their clients greater customization and better products and services.

While these partnerships can deliver better services and products to clients, there have been concerns around the issues related to data sharing, such as data privacy and security. The same principles that are applied internally around transparency and accountability should apply to third parties.

Ensure third-party confidentiality

It is vital that privacy policies inform customers in a concise, intelligible, and easily accessible way how and what data is being shared with third parties not acting on behalf of the FI, the reasons for doing

[24] From mystery to mastery: unlocking the business value of Artificial Intelligence in the insurance industry, Deloitte, page 32, https://www2.deloitte.com/content/dam/Deloitte/ru/Documents/financial-services/artificial-intelligence-in-insurance.pdf.

so, and the steps involved in protecting that data. Where consent is revoked, financial institutions should explain the implications and how it may affect the customer's experience.

In simple terms, third party transparency translates into ensuring that shared private and sensitive data are treated confidentially, and no traces of identity are exposed. Thus, FIs need to make sure that partners have restrictions on whether and how private information can be distributed further.[25]

Beyond compliance the focus should be in maintaining customer trust. Apart from regulatory or operational requirements, FIs should apply the same principles to third parties as those applied internally.

Principles around knowledge sharing also come into play - as customers become more aware of the value of their data in terms of benefits and risks of sharing data, they can make better informed decisions.

Customers should be able to easily request and obtain information on whom their data is being shared with, and where possible, how their data is being shared and processed by third parties not acting on behalf of the FI. Moreover, in the event of a cyber breach of a third party, FIs should notify (and depending on the jurisdictions, are compelled to inform) any impacted clients as soon as possible and be able to identify and explain to their customers where and when their data was improperly used or accessed.

Thorough review of partners accessing customer data

Financial institutions need to conduct a proportionate risk assessment and thorough review (due diligence) of their potential partners and their data protection and security processes before engaging and during their relationship. This would ensure that, in case they have access to their client's information, those data are as safe with the service provider.

FIs need to have confidence in the security of the data, both in terms of sharing it with partners and the way it is stored. In practice, this means that FIs should make sure that partners accessing customer data have adequate data protection and security processes and standards[26] in place.

Ensure data privacy through enforceable contracts with partners

Financial institutions should aim to make sure that partners with whom they share their customers' data are bound by an enforceable agreement which outlines how that data can be used.

Principles at Play: *BBVA Mexico uses big data to boost tourism in Mexico*[27]

BBVA Mexico and the Mexico's Ministry of Tourism (SECTUR) have presented a pioneering big data application that identifies trends and productivity in the country's tourism destinations. This case illustrates how using anonymous and aggregated customers' data can be very beneficial in boosting a country's economy.

During the first stage, the project gathered information on the trends registered by 86 million national and foreign bank cardholders over a one-year period in 12 statistical areas that cover the 111 Magic Towns of Mexico and the main tourist areas in the country. Of these, 20 million are BBVA Mexico customers. In one year, 1.5 billion transactions were made, totaling 813 billion Mexican pesos. The project was developed with

[25] "5 Principles for Big Data Ethics," Medium, September 14, 2018, available at https://towardsdatascience.com/5-principles-for-big-data-ethics-b5df1d105cd3.
[26] An example of internationally recognized information security management standards is the ISO/IEC 27000-series published jointly by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC).
[27] https://www.bbvadata.com/es/big-data-to-enhance-the-tourism-in-mexico/

strict adherence to data protection and the privacy rights of financial service users, and only statistical insights were shared.

The massive adoption of *big data* technology means that cardholders now generate regular information about payments that can be analyzed. BBVA Mexico, BBVA Data & Analytics, and SECTUR then collaborated to extract and analyze massive amounts of data to understand the activity in a given area. This analysis, which was done by studying the anonymous digital footprint of the visitors, allows the bank to know, for example, that Cancun mainly attracts visitors from the United States and Argentina, while in Cozumel, Mexicans mostly spend money on restaurants and travel.

While the high-resolution data came from BBVA Mexico, the analytical tools, the results, and the statistical and geographical visualizations were provided by BBVA Data & Analytics. Mexico's Ministry of Tourism hired experts to interpret the statistics and prepare the study's findings. Some of the main conclusions can help tourism companies to offer better service during peak spending hours, which is not easy to do through traditional surveys. For example, they can anticipate the movements made by visitors to islands and the spending distribution according to the time of day.

This study once again shows that monitoring tourists' actual spending is a useful tool for shaping a territory's production structure and the changes and trends that affect the tourism sector, with the goal of implementing policies that promote and improve the country's tourist destinations.

Principles at Play: *Banco de Crédito del Perú (BCP) prioritizes customer safety through enforceable contracts*

When partnering with a third party, BCP guarantees the safety of their customer's data by signing an enforceable contract.

Typically, third parties do not receive Personal Identifiable Information (PII) nor highly sensitive customer data. However, when it becomes necessary to share such data, BCP performs due diligence of the partner, sends data in an encrypted format, and establishes a process that guarantees that BCP can track any shared data by inserting a tracking code into every customer data that is shared. Inserting seed names helps BCP identify any data leakages, when and where a breach occurred and can be able to provide customers with adequate and timely information.

Establish accountability through compensation and dispute settlement frameworks

Financial institutions should define accountability and ensure it is understood by all third parties they collaborate with. As part of this, dispute settlement frameworks should be established that clearly articulate how to resolve disputes between them, and how to assign liability in the event of incidents or errors.

A system (dispute settlement framework) should be put into place to ensure fair compensation of customers in the event of an incident or errors arising from a third-party partner access to client data that leads to damages. Examples include stolen identity, unauthorized transactions, or defective payments.

FIs should provide appropriate avenues for customers to challenge or dispute data accuracy or data use, based on risk or harm to the individual, and make provision for the third-party vendors to assume appropriate liability. Accountability should be considered through the lens of data ethics, and not only through that of compliance, as this will aid in ensuring the ethical treatment and sharing of customer data.

5. Skills, Awareness, and Knowledge Sharing

Training and education on the ethical issues around the use of data is an essential element of the DEC. At its core this involves actively and regularly training employees on data ethics and the requirements to support an ethical framework. Similarly, this includes sharing knowledge with customers about the evolving landscape of data usage, and helping customers better manage their data and make well-informed financial decisions.

Apart from training and awareness, the organization should have a culture of doing the right thing, and ethical decision making should be embedded in every day processes as opposed to a tick-box exercise.

Enhance digital awareness and literacy for customers

Financial institutions—possibly in cooperation with the public sector—should invest in transmitting knowledge that allows customers to better understand how data-based technology and business models work and assess risks and benefits.

In a financial services context, it involves a range of initiatives, such as advisors who take additional time to explain the pros and cons of products, webpages updated with information customized to the client's needs or the use of interactive online capabilities and platforms (e.g., chats, virtual personal assistants). Such initiatives have the effect of increasing customers' digital awareness and their knowledge of both the company and the market – ensuring that the customer understands what is being done with their data, how consent mechanisms work, and the risks and benefits of new products and services. At its core - helping customers get the best use out of available products and services.

The benefits of knowledge sharing are shared by all vital players in the financial ecosystem (i.e., financial institutions, third parties, and regulators) as customers can make better informed decisions and identify ways in which they can save money.

Increase customer awareness coupled with clear ethical data practices can establish trust as a key differentiator, while also informing (e.g., by displaying notices and guidelines), educating (e.g., advisors explaining the pros and cons of products), and empowering (e.g., giving customers reliable options for making informed decisions about how their data is used) customers.

Enhance data training and education for employees

Financial institutions should regularly provide employees with firm-specific internal data policies and processes to ensure that staff are knowledgeable and capable of following complex and evolving data practices.

Actively and regularly training all employees on the definition of data ethics, and on the requirements to support an ethical framework is essential to protecting customers. All employees should receive such training. Beyond that, different training and awareness interventions should vary based upon an employee's role within the institution (e.g., branch teller vs. product developer vs. analyst). For example, depending on the employee's position within the firm, training on testing and evaluation of outcomes of analytics may be needed. Additionally, this includes how the data is handled, and ensuring that people with the right skills and experience are developing models using new analytics. We discuss challenges around algorithmic decision-making systems in more detail below.

Financial institutions' employees play an important role in implementing procedures and regulatory requirements related to the use, protection, and overall management of customer data. Regular and thorough

data ethics trainings to employees who handle customer data—regardless of their position—is needed to provide staff with clear guidance and to keep them updated on examples of good practices.

Furthermore, FIs need to regularly align their internal policies and regulations and make them available to staff in an easily understandable way and with clear instructions and deadlines on use and implementation. This requires ensuring that skill requirements are met by applicable employees so that they can follow and use complex and evolving data practices and provide customers with correct and adequate information. In addition, staff evaluations should consider employee's adherence and compliance to the firm's internal data policies to uphold high standards for the FI.

A solid and deep understanding of company-wide data ethics by FI employees is crucial and a prerequisite to upholding the various ethical data principles outlined below, as is the need for each financial institution to develop a definition of fairness at a strategic level. More on the latter is described in the next section.

Principles at Play: *BBVA's innovative training programs and partnerships*

The availability of talent that understands how to handle data to solve business issues is scarce. BBVA, a multinational bank headquartered in Spain, has tackled this issue by developing innovative training programs for existing employees and partnering with universities to attract (and train) talent, as well as other recruitment efforts. These efforts allow BBVA to ensure that their goal of using innovation to transform its business and bring the age of opportunity to everyone, including its customers, is accomplished.[28]

Their innovative training program was launched in 2015 and since then has retrained hundreds of employees preparing them to find patterns and use computer programming skills with new processing and visualization tools. The training program focused on upscaling skills of current employees building up in-house knowledge on new tools and techniques. This was done initially with the creation of the training course "From Data Mining to Data Scientist",[29] but has now evolved towards a much bigger initiative called "Data University". Through its innovative training program, BBVA's employees learn the importance of data in meeting the challenges of day-to-day business, improving processes, and even creating new products and services.

Additionally, BBVA has worked together with the Universidad de Navarra to tailor-made certificate, Master's and doctoral level programs that can potentially train their future pool of applicants and support data science research.[30] This partnership allows BBVA to attract talent among Universidad de Navarra's graduates, ensuring that they acquire the appropriate data scientist skills and ethical values needed to understand customer's data privacy.

[28] This was the core message the BBVA CEO delivered in a keynote speech to a packed audience at the 2017 Money 20/20 event in Copenhagen.

[29] https://www.bbva.com/en/bbva-can-also-data-scientist/

[30] https://www.bbva.com/en/bbva-and-universidad-de-navarra-work-together-to-train-their-employees-and-support-data-science-research/

Appendix A: Further Considerations

1. Responsible Data Management Cycle

The compliance aspect of providing financial services includes the lawful collection[31], quality assessment, processing, safeguarding, analysis, and destruction of data elements that come from various sources and span across a broad scale, including financial assets and liabilities, as well as sensitive information about customers.

Having a sound data governance and management framework (of policies, procedures, controls, and good practices) has always been considered a cornerstone of financial institutions' operations, with particular emphasis on data protection and security against losses or breaches and preventing the internal misuse of data. Many implemented modalities are rooted in some form of regulation but are also complemented with additional safeguards adapted to each financial institution's operations.

In general, financial institutions publish data privacy notices and policies that disclose what information they capture, the purpose of data processing, and whether data are transferred to third parties within or beyond the institution.

Legitimate use

FIs should have a lawful basis for the processing of any data, and technological solutions should confirm data is only used when a legal basis exists.

2. Data Control

Digital products should empower customers by helping them make informed decisions about their privacy and give them easier control over their data.

Data management transparency

Frequently, it is the financial institutions' original intent to provide a clear and short statement around data collection and privacy, but when the legal requirements are included, it can become long and unwieldly. Financial institutions should minimize unnecessarily complex and lengthy language by eliminating jargon, and proactively improve the digital literacy and awareness of their customers. Privacy policies can be redesigned to provide clear and accessible information regarding a firm's data processing methods, where appropriate. This may include clearly labeling sections, creating expandable text that allows customers to locate the information they need, and by giving contextual explanations that make legal language easier to understand.

Data mobility

In jurisdictions where Open Banking regimes are compulsory, customers willing to share data with third parties or to exercise their right to port their data to another provider should expect financial services companies to provide access/permission to these third parties. As a general principle, any data sharing framework should be reciprocal in terms of access to data (i.e., the entities obliged to make their customer's data shareable should also be able to gain access to data from third parties, should the customer request it), while respecting the data sharing regulatory framework in the respective jurisdiction.

Furthermore, there can be specific cases when data sharing becomes necessary across institutions for societal benefits, for example, for financial crime detection. It is important that in such cases the need to protect individual data is carefully balanced relative to the broader societal benefit.

31 In addition to acquiring customer consent, financial institutions should ensure that the amount of data being collected is proportional to the purpose for which it is being collected.

3. Challenges Around Algorithmic Decision-Making Systems

Algorithmic decision-making systems, such as more automated credit scoring systems were established over a decade ago to help businesses "reduce the possibility of unlawful discrimination by helping ensure consistency and uniformity while minimizing individual judgement and discretion."[32]

In recent years there have been several examples of the opportunities of leveraging existing data for financial health and combatting fraud, among others. According to a study by Adobe Analytics, consumers are open to embracing new technologies in banking with 44% of Generation Z and 31% of Millennials having interacted with a chatbot, with the majority of those that did, preferring it over interacting with a human representative.[33]

In credit markets, a combination of alternative data and ML has helped make financial services more inclusive and accessible to customers. A 2019 FINREG study shows that using variables such as rental payments and utility payment histories can serve as reliable markers to determine a customer's ability to repay. [34]

Additionally, a UC Berkeley study shows that alternative data has substantially decreased pricing disparities and eliminated underwriting discrimination for Black/Hispanic borrowers.[35]

Bias and discrimination

Definitions for bias are various and can be viewed under many different lenses. There is of course the statistical definition being systematic differences between a population subsample and the population at large, and the more popular or social definition being a human judgement based on preconceived notions. We define "bias" as an unfair inclination for or prejudice against a person, group, object, or position.

Discrimination, whether intentional or unintentional, may occur when one group of people is more adversely affected by a decision or process than another group without a legitimate and neutral justification.[36] This area continues to evolve, along with the supporting literature. Further comments on discrimination metrics as defined by the fairness literature include equalized odds and equal opportunity[37], or the WAE-WYSIWYG spectrum[38], each of which reflect different worldviews depending on the context of the problem domain.

The IIF Thematic Paper titled "*Bias and Ethical Implications in Machine Learning*[39]" delved into this topic more closely. Therein, we defined ethics as a system of moral principles governing a person's behavior or conduct of an activity. Relating to financial institutions, regulation says what FIs can do, while ethics says what they should do.

Establish appropriate AI/ML governance

Humans translate an initial task into constraints and objectives; thus, the effectiveness and reliability of an algorithm depends on the human that delineates the task, among other factors such as the features used to train the model. One scenario in the medium term is that as ML algorithms change their behavior (i.e., results) with feedback data, the results will start to vary based on new information that is continually included

---

[32] FDIC Supervisory Insights, "*Fair Lending Implications of Credit Scoring Systems*" (Summer 2005)
[33] Adobe Analytics, "*How Different Generations Bank*" (2019). Available at: https://theblog.adobe.com/adobe-analytics-research-how-different-generations-bank/
[34] FinReg Lab, "*The Use of Cash-Flow Data in Underwriting Credit*" (2019). Available at: https://finreglab.org/wpcontent/uploads/2019/07/FRL_Research-Report_Final.pdf
[35] Bartlett, Robert et. al, "*Consumer Lending Discrimination in the Fintech Era*" (2019). Available at: https://faculty.haas.berkeley.edu/morse/research/papers/discrim.pdf
[36] Laws typically evaluate the discrimination using two distinct notions: disparate treatment, and disparate impact. Disparate treatment includes overt discrimination, as well as more subtle unjustified differences in outcome on a prohibited basis. Whereas disparate impact occurs when a neutral policy or practice results in a disproportional exclusion or burden on certain group of people, whether the policy was created with the intent to discriminate.
[37] *Equality of Opportunity in Supervised Learning*, Hardt et al. 2016
[38] *On the (im)possibility of fairness*, Venkatasubramanian et al. 2016, accessed at: https://arxiv.org/abs/1609.07236
[39] IIF, *Machine Learning Thematic Series: Bias and Ethical Implications in Machine Learning*, May 2019, https://www.iif.com/Publications/ID/3338/IIF-Paper-on-Bias-and-Ethical-Implications-in-Machine-Learning

in each iteration or run of the analysis. The current reality, for the most part, is that the frequency of model recalibration with feedback data is not so high, and for most models, it is far from being upon real time data.

Establish accountability and define fairness

We identified four main factors related to data and algorithmic decision-making solutions that need to work together to generate optimal results: *the volume of data, the quality of data, the diversity and representativeness of system engineers and data scientists,* and *the outcome of the algorithmic decision-making system*. These factors all have the principle of accountability in common and are described below.

In terms of the *volume of data* – algorithmic decision-making systems that employ AI/ML allow for greater use of data, resulting in more reliable results and potentially improving the fairness of the decisions.[40] To give a few examples, ML is currently being used in financial services to automate operational processes, enhance security, develop more accurate pricing models, automate marketing campaigns to enhance customer relationships, and make recommendations on training of internal teams.

In terms of *quality of data* – data collected by any human-designed process will have some bias in it, and historical datasets will always reflect historical biases. Data used to train a system can underrepresent or overrepresent members of certain protected classes or carry over past discriminatory practices, whether intentionally and caused by prejudice, or unintentionally and caused by limitations in knowledge and experiences. Data quality of inputs is important, but those inputs are being chosen because of a given desired output, therefore it is not enough to look at the input and the model.

In terms of *outcomes of the algorithmic decision-making system* - trying to mitigate biases by restricting inputs to models should not be the only approach. Alternative approaches should be considered, such as looking at the outputs and outcomes of those systems.

There is an additional differentiation between outputs and outcomes. To illustrate, an output will be a prediction that someone with a mental condition will commit a crime, the intervention or outcome will be the decision that a human makes on how to act on that prediction. One response may be to arrest, while another very different response may be to send the person to rehabilitation. Accordingly, how the output of the algorithmic decision-making system is utilized is an important area where potential bias can be mitigated.

Sensitive attributes

"Sensitive attributes" are legally protected classes or sensitive features—such as race, gender, disability, or religion—that could create unjust or harmful outcomes (i.e., moral, ethical, and legal problems) when used in a model. These are prescribed in data privacy laws.

Firms need to be careful how they address unfair bias in their models. Mitigating biases by restricting inputs to models should not be the only approach taken, alternative approaches, such as looking at the outcomes of those systems should also be considered.

Existing restrictions on the use of sensitive personal information make it more difficult for firms to determine if an algorithm discriminates based on a protected characteristic. In fact, systems can produce a disparate impact due to the correlations between the variable within a sensitive/protective class and other closely correlated variables, for example, zip code and race.

There is no consensus on the best way to deal with sensitive attributes. For many the complete removal of sensitive attributes is ineffective since they are necessary to produce accurate outputs and may be inferred from non-sensitive attributes. The practice/control mechanism of excluding sensitive attributes from the beginning and not including them as part of the feature analysis/selection/engineering process varies by region.

[40] More volume of data can also make it easier to find and confuse spurious relationships with valid relationships, which highlights the need for robust model development measures. For information on explainability and interpretability of these algorithmic decision-making systems, see our IIF report, *Explainability in Predictive Modeling,* November 2018, https://www.iif.com/Publications/ID/1423/Machine-Learning-Paper-on-Explainability-in-Predictive-Modeling.

Non-discrimination laws and data protection laws demand ethics and fairness, and in many cases prevent people from being discriminated against on the basis of certain protected characteristics. In the U.S., firms are subject to fair lending statutes that prohibit discrimination in lending,[41] among other standards, in addition to firms' own internal codes of conduct. The EU's General Data Protection Regulation includes "special categories" of data and strengthens individuals' rights in and control over their personal data.[42]

In the U.S., fair lending statutes that prohibit discrimination in lending predates ML. In the U.S. firms must not use prohibited basis data or proxies for discrimination, and the direct utilization of sensitive attributes in model production or development is viewed by many as unlawful under current U.S. anti-discrimination laws. Additionally, as part of the risk assessment procedures, such attributes or anything acting as a proxy for any protected class are excluded before development commences.

Similarly, in the EU data protection laws have long included "special categories" of data. Although, this view of whether to include or exclude sensitive attributes is very much linked to whether it is lawful or unlawful to do so. For some executives, withholding sensitive demographic information from an algorithm does not solve the problem of "redundant encodings," where membership in a protected/sensitive class is encoded in other data. Therefore, if the training dataset is rich in features that are granular and diverse, the algorithm given particular pieces of data and values that are highly correlated can then deduce a certain protected/sensitive characteristic, even if implicitly.

Ensure algorithmic decision-making transparency and explainability

Firms should ensure that algorithmic decision-making solutions respect and integrate "elements of trust"[43] by performing an impact assessment of use cases considering the potential impact on business continuity and the customer's fundamental rights and access to services.

To illustrate, in the financial context, a single ML model may have multiple stakeholders with their own unique use cases. Thus, the controls should be commensurate with the impact of each specific use case. One single ML model may have several different types of stakeholders, for instance: those implementing an ML application, management responsible for the application, the FI's independent control functions, conduct regulators, and prudential regulators. Thus, the impact of an autonomous decision, such as whether a loan gets processed, would be of interest to not only those tasked with implementing the ML application (to understand outliers), but also to a conduct regulator, a data protection supervisory authority, and a prospective customer. However, other stakeholders may be more interested in how the model works more generally.

Clearly defined human controls

Human-in-the-loop

Human-in-the-loop (HITL) is the process where decisions made by the ML application are only executed after review or approval from a human. This process starts by involving human intervention in training stages when building an algorithm, creating a continuous feedback loop that allows the algorithm to give better results, and for testing and validating the model.

Results and improvements deriving from these controls need to be fed back to address concerns. This means that it should be possible to amend which information is selected, or to build a system that allows to interfere with the calculations of an algorithm by a human (in practice this is currently the case).

---

[41] *The Fair Housing Act* (FHA) and *Equal Credit Opportunity Act* (ECOA) protect customers by prohibiting unfair and discriminatory practices. ECOA prohibits discrimination on nine bases – race, color, religion, sex, national origin, age, marital status, receipt of public assistance, or exercised any right under the Consumer Credit Protection Act. Regulation B implements ECOA. FHA covers residential real-estate transactions, prohibits discrimination on seven bases – color, race, religion, sex, national origin, handicap, and familial status.

[42] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016

[43] EBA, *Big Data and Advanced Analytics*, January 2020, remarked on the need to include "elements of trust": https://eba.europa.eu/sites/default/documents/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf.

HITL is used by firms for various reasons as a combination of human and machine intelligence to create a continuous loop for training, testing, finetuning, validating, and monitoring ML algorithms. In practice, HITL is integrated by humans labeling the training data, which is later fed to algorithms, or/and by humans during model validation where they check and evaluate ML predictions. Where results are inaccurate, data is re-checked and fed back into the algorithm.

The HITL approach is used for different types of data labeling from labeling structured data (e.g., transactional data) to semi and unstructured data (e.g., language or voice recognition, video, images, and other content that is not clearly labeled). HITL is therefore used to make such content comprehensible.

Based on our discussions, we see HITL increasingly being used as a safeguard. Firms' use of HITL highlight that it is important to put the human in the center of ML models, and that the ML should be designed to augment human intelligence, not replace it. When we apply this thinking to the financial industry, this translates to risk appetite. Therefore, even if a FI decides to use ML for a particular application, there should be an ongoing assessment of how ML should evolve based on the financial industry's risk appetite. Many firms remarked that the use of HITL goes hand in hand with the design principle of the ML model.

Additional design solutions

- Address disparities before putting model into use.
- Define own approach to algorithmic decision making and bias mitigation (i.e., rules, criteria, red lines, explainability efforts over non-interpretable models, etc.), document them and make them available within the organization.
- Given that proxies can be very difficult to find, a process should be developed to help identify proxies.
- Assign a lead ethics official that oversees how algorithmic decision-making systems arrive at their results and testing for bias in accordance with the approach previously defined.
- Generate new data to improve the accuracy of the model. New data can then be used to create more adequate datasets. Its creation can be costly and may not have a private-sector payoff. In practice this would mean for instance, extending credit to members of a protected class who might traditionally miss eligibility requirements, and then analyzing the results to identify creditworthiness. This type of exploration rarely takes place because it would require granting some loans randomly without regard to the properties of applicants.
- Training programs to support these regulatory agencies in their expanding roles.

4. Partnerships and Trusted Third Parties

The regulatory community has responded, albeit in an inconsistent manner, resulting in a growing patchwork of data sharing regulation around the world. Herein, we highlight ethical guidelines and standards that FIs worldwide should, at a minimum, strive to uphold when data sharing and privacy regulation is limited, weak, or non-existent.

Ensure data privacy through enforceable contracts with partners

In the case of international partnerships, there should be a governing law clause in the agreement that clearly defines that the contract abides by a particular jurisdiction. In the event of a disagreement, the contract would indicate which country's court system would be used to resolve a dispute.

Establish accountability through compensation and dispute settlement frameworks

Defining accountability and ensuring it is understood by all third parties is paramount as this will help resolve ambiguities involving liability should they arise.

Bilateral agreements should include requirements for appropriate insurance or funds for such instances. As per a 2018 BaFin report on big data and artificial intelligence, a customer "can only make a sovereign decision

if they are adequately informed about the potential reach and consequences of the use of their data, if they are given reliable options for controlling how their data is used, and if they have actual freedom of choice."[44]

5. Skills, Awareness, and Knowledge Sharing

Enhancing digital awareness and literacy for customers

Though many customers are comfortable sharing their personal information in exchange for services, they may still be wary about its specific use and purpose. Not only is this essential to maintain and enhance customer trust, it is also a prerequisite for the successful promotion of digital literacy and the development of the financial ecosystem.[45]

Because expertise is built slowly over time, efforts to reduce the asymmetry in knowledge between financial institutions and their customers around data management and financial products is undoubtedly a long-term investment, but also a great opportunity that can have a strong and positive impact on all parties involved.[46] For example, this type of knowledge sharing can facilitate the engagement of customers in new ways.

Finally, because digitization is driving the democratization of finance and deepening inclusion, FIs need to be cognizant of how to engage and share knowledge with their new, and previously unbanked, customers.

[44] BaFin, "*Big data meets artificial intelligence*" (July 16, 2018), can be accessed at: https://www.bafin.de/SharedDocs/Downloads/EN/dl_bdai_studie_en.html.
[45] Liability and consumer protection in open banking, IIF, September 2018: https://www.iif.com/Publications/ID/1418/Liability-and-Consumer-Protection-in-Open-Banking
[46] Customer education increases trust, MIT Sloan Management Review, October 1, 2008: https://sloanreview.mit.edu/article/customer-education-increases-trust/

Appendix B: Glossary

This section aims at clarifying the definitions used in this document to facilitate easy understanding of key terms and expressions.[47]

Algorithm:
An algorithm is an unambiguous procedure to solve a problem or a class of problems. It is typically composed of a set of mathematical instructions or rules given to a computer (or a network of computers) that take some input data and return outputs. Algorithms can be combined to develop more complex systems, such as web services or autonomous cars. An algorithm can be hand-coded, by a programmer, or generated automatically from data, as in machine learning.

Algorithmic decision-making (ADM):
ADM is a specific type of algorithm aimed at supporting business processes and decisions. Algorithmic or automated decision-making occurs when the computer determines the most optimal decision based on previous decision data and a multitude of inputs. It may assume varying degrees of human involvement. Semi-automatic ADM assists humans in making decisions by suggesting an optimal decision, but the human makes the ultimate determination based on a given suggestion. For example, an ADM can assist doctors in identifying diseases and help them to make diagnoses. An ADM can also be used to take fully automated decisions, as in automated metro systems. Very often, they are used to make predictions or to estimate risks. Table A provides some examples of ADM applications:

Table A: Examples of applications of ADM by objectives and types of users[48]

Objectives / Users	Individuals	Private sector	Public sector
Improvement of general knowledge	N/A	Drugs discovery	Climate Weather forecast Environment Healthcare
Digital services	Quantified-self Finance Note taking Smart home Recommendations	Risk scoring Payment systems Targeting Personalized services	Predictive justice Predictive policing Hazard prediction Infrastructure development planning
Physical systems	Autonomous cars Home robots Security Personal assistants in the home	Autonomous robots	Autonomous weapons Defense Transport Smart cities Smart grids

Algorithmic interpretability:
We say that a model is interpretable when its internal operations can easily be understood by a human due to their low level of complexity. All rules-based algorithms and most less sophisticated ML algorithms (such as regressions) are interpretable.

Algorithmic explainability:
Most sophisticated ML algorithms (such as DL or CNN) are not interpretable[49], however, ex-post extra development efforts can be carried out to provide explanations about how factors in the input led

[47] The content related to key terms of algorithmic decision-making builds on the following document: European Parliament, *Understanding algorithmic decision-making: Opportunities and challenges*, March 2019, https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf.
[48] Source: European Parliament, *Understanding algorithmic decision-making: Opportunities and challenges*, March 2019, https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf.
[49] A taxonomy of ML techniques and their degree of interpretability can be found here: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence/annexe-2-algorithmic-techniques/

to the outcome, without getting to know the intrinsic operations of the model. Fairness risks mitigation efforts require the application of those explainability techniques.

Artificial intelligence (AI):
AI is the capacity for machines to resemble human intellectual abilities. Narrow, or weak, AI is designed to perform a specific task, such as facial recognition or product recommendation. General, or strong, AI aims at outperforming humans across multiple domains.

Consent:
Consent is defined as any freely given, specific, informed, and unambiguous indication of the data subject, i.e., the customer's wishes by which the customer, by a statement or by a clear affirmative action, signifies agreement to the processing of data relating to them.

Customer data:
Customer data refers to both the qualitative and quantitative forms of information that is collected and/or created during interactions between a provider of a good or service and the individual customer of a product or service (e.g., a financial institution and its customer). It can include—but is not limited to—personal, behavioral, and demographic data. Moreover, it can be provided (e.g., a customer filling out a mortgage application form), observed (e.g., a customer's loan payment history), inferred (e.g., a bank's underwriting models determining the loan size a customer qualifies for) or obtained from third parties (e.g., FICO score).[50] It is, therefore, much broader than "*personal*" or "*personal identifiable data (PII)*."

Table B highlights some of the most important traditional and non-traditional types of customer data and genuinely reflects the changing nature of data availability from paper-based forms to real-time elements and across all aspects of life:

Table B: Traditional and Non-Traditional Types of Customer Data[51]

	Traditional forms	Non-traditional forms
Financial	Bank statements, credit scores	Peer-to-peer payments, online budgeting, mobile payment applications
Health	Medical records, insurance claims	Fitness tracking, sleeping/eating habits
Identity	Public records, tax filings	Biometric data (e.g., fingerprints, photos, iris scans, face recognition programs), digital IDs
Location	Telephone books	Geolocation tracking
Media behavior	Library checkout histories	Web browsing activities, content streaming
Social	Organization registries	Social media connections and activities

Ethics:
Ethics is defined as a system of principles governing a person's behavior or the conduct of an activity. In the case of the financial institutions, ethics bridges the gap between regulated and non-regulated spaces - that is, firms know what they can do in accordance with relevant laws and regulations, but ethics guides firms on what they should do (what is right or wrong). Financial institutions have long established ethical standards that are enshrined in firms' values and codes of conduct, incremental to those that are adopted in response to regulatory requirements such as those relating to fair lending or best interest standards. It is important to note that what is deemed "ethical" varies between individuals, societies, and jurisdictions, and can change over time.

[50] "The Appropriate Use of Data in Financial Services," World Economic Forum, September 2018, p.13, http://www3.weforum.org/docs/WP_Roadmap_Appropriate_Use_Customer_Data.pdf.
[51] World Economic Forum, "*The Appropriate Use of Data in Financial Services*" (September 2018), Figure 10, p.13

Fairness:
Fairness—a hotly debated term in the global public policy discourse—is defined here as a principle that understands and minimizes discrimination based on legal protected characteristics that aligns with a firms' ethical values and considers the dynamics and cultural variations of ethical codes.[52]

Financial institutions (FIs):
Financial institutions are those that predate the digital revolution (e.g., traditional banks, insurers, asset managers). They are considered the incumbents in the industry, as opposed to the new market entrants such as financial technology firms (commonly referred to as "FinTechs") and larger more established technology firms (commonly referred to as "Bigtechs") that also offer financial products and services.

Machine learning (ML):
ML is an AI component that provides systems with the ability to automatically learn over time, generally from large quantities of data. The learning process is based on observations or data, in order to identify patterns in data and make better predictions. An ML algorithm can therefore be an algorithm that, from data, generates another algorithm, usually referred to as a model. For example, the Amazon recommendation algorithm uses customers' profiles to learn which products are likely to be of interest to them. When users visit the Amazon site, the recommendation model built by the system uses their profiles to produce personalized recommendations.

Personal data or personal identifiable information (PII):
Personal data, or personal identifiable information (PII), is defined as any data that, when used alone or with other relevant data, identifies or can identify an individual (e.g., name, address, social security number, etc.).

Sensitive personal data:
Personal data which are, by their nature, sensitive in relation to fundamental rights and freedoms. For example, racial or ethnic origin, political opinions, religious or philosophical beliefs, sexual preferences, biometric and genetic, etc.

Third Parties:
An organization that has entered business relationships or contracts with a firm to provide a product or service.

[52] We believe that prior to implementing techniques and guidelines related to bias mitigation, FIs must build an internal definition of fairness at a strategic level. Data scientist cannot be accountable for the implementation of firewalls against discriminative outcomes in their models if fairness has not been defined at a strategic level; until then they work under uncertainty.

Lead Authors

Natalia Bailey
Policy Advisor, Digital Finance
nbailey@iif.com

Dennis Ferenzy
Associate Economist, Digital Finance
dferenzy@iif.com

Other Contributor

Brad Carr
Managing Director, Digital Finance
bcarr@iif.com