

Credits

This report was prepared at the request and with the financial support of the Autorité des marchés financiers by **Algora Lab** (Université de Montréal) with the support of the **International Observatory on the Societal Impacts of AI and Digital Technology** and the scientific contribution of **CIRANO** (Polytechnique Montréal).

The report was drafted under the direction of **Marc-Antoine Dilhac** (prof. Université de Montréal, dir. Algora Lab, CIFAR-Canada AI Chair), with contributions by **Rhéia Khalaf** (Université de Montréal, Fin-ML), **Nathalie de Marcellis-Warin** (prof. Polytechnique Montréal, dir. CIRANO), **Manuel Morales** (prof. Université de Montréal, dir. Fin-ML), **Robert Normand** (CIRANO, Polytechnique Montréal) and **Gabrielle Veilleux-Verreault** (Université de Montréal).

The interviews with stakeholders were organized by Robert Normand and conducted with Marc-Antoine Dilhac, Gabrielle Veilleux-Verreault and Manuel Morales.

The Algora Lab team (Marc-Antoine Dilhac, Gabrielle Veilleux-Verreault, Carl Maria Mörch, Fatima Salazar Gomez, Camylle Lanteigne, Lucia Flores Echaiz, Vincent Mai) prepared and organized the deliberation workshops on artificial intelligence in finance with the Authority. Marc-Antoine Dilhac and Gabrielle Veilleux-Verreault designed the deliberation scenarios and analyzed the contributions. Carl Maria Mörch designed the first steps of the project and prepared the scientific documentation.

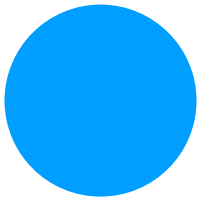
We thank the Authority for its support in organizing the deliberative workshops and for the dialogue we had with the steering committee. We are particularly grateful to Lise Estelle Brault, whose patient review of the draft report was invaluable, and to Oumar Diallo and Kim Lachapelle. We also thank Hélène Samson, Jean-François Ouellet and Benoit Vaillancourt.

Partners :



Centre interuniversitaire de recherche en analyse des organisations





Foreword



Louis Morisset

President and CEO

Autorité des marchés financiers

The accelerating digital transformation in the financial services industry is resulting in major changes in the processes of many businesses. Technological innovations such as advanced analytics and artificial intelligence (AI) are being deployed at a rapid pace and are contributing to an increase in the number of options available to consumers. Moreover, access to big data on consumer profiles and patterns means it is now possible to pinpoint the needs of consumers more precisely and to offer them personalized financial products and services. Although this profound

transformation is yielding tangible benefits, it is also bringing to light significant issues requiring our consideration.

In its 2021–2025 Strategic Plan, the AMF set an objective of providing strong thought and action leadership on current and emerging issues. We intend to propose solutions to such emerging issues and to develop, together with all stakeholders, standards that promote both innovation and consumer protection. Consequently, the AMF is publishing 10 recommendations today to help guide financial industry participants, and the AMF, in the responsible use of AI. The AMF is proud to be the first Canadian regulator to contribute to this important discussion.

For this project, the AMF was supported by a research team under the direction of Marc-Antoine Dilhac, Associate Professor of Ethics and Political Philosophy at the Université de Montréal and Director of Algora Lab, an interdisciplinary laboratory advancing research on the ethics of artificial intelligence. Professor Dilhac launched and had a major hand in developing the *Montreal Declaration for a Responsible Development of Artificial Intelligence* in 2017, a charter that has become a worldwide reference.

The AMF also received input from financial consumers who agreed to share their perceptions of the issues raised by the use of AI in various scenarios related to the financial markets. Like the work that led to the Montreal Declaration, the recommendations in this report are based not only on the findings of experts in the field but also on discussions with citizens who are required to interact with new technologies on a daily basis.

I wish to sincerely thank Professor Dilhac and his team, as well as the consumers who took part in the workshops, for their contribution to this project.

The publication of this report is intended as a starting point for dialogue among financial industry participants, financial consumers and the AMF about the responsible use of AI. However, I strongly encourage financial sector participants to consider these recommendations going forward when developing AI systems. The AMF pledges to do likewise as it continues its own digital transformation.

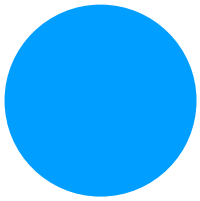
Montréal is known today as a global hub of AI development. This is a source of pride for us. Let's continue to cultivate this local expertise in our financial sector businesses so that we might derive maximum benefit from this new technology. However, let's do so responsibly, while respecting financial consumers, so that everyone wins.



Sommaire

Credits.....	2
Foreword	Erreur! Signet non défini.
Executive summary	Erreur! Signet non défini.
10 Key recommendations	13
Introduction.....	15
1. AI in financial services and investment.....	17
1.1. Background and definitions	17
1.2. Using AI in finance and investment: taxonomy by function	20
1.2.1. Assessment.....	21
1.2.2. Nudge	23
1.2.3. Optimization.....	26
1.2.4. Advice and information.....	28
2. Ethical risks and challenges of responsible AI deployment in finance.....	30
2.1. The perspective of financial institutions.....	31
2.1.1. Privacy	31
2.1.2. Equity.....	32
2.2. Consumer expectations	32
2.2.1. Respect for autonomy	33
2.2.2. The risk of surveillance.....	34
2.2.3. Equity and solidarity.....	34
2.2.4. Responsibility	35
3. Assessing the goals of AI in finance upfront.....	36
3.1. The relevance principle.....	36
3.2. The resilience principle	38
3.3. AI and the development of responsible and sustainable finance (ESG)	39

4. Ethical compliance principles	41
4.1. Frameworks for the framing of AI in the financial sector	43
4.2. An encompassing framework for AI ethics in finance	46
4.3. Equity and discrimination	49
4.4. Transparency, explainability and justifiability	51
4.5. Autonomy and nudging AIS.....	53
4.6. Solidarity.....	55
5. AI governance mechanisms	57
5.1. Ethical conduct.....	57
5.2. Impact assessment and audit	58
5.3. Certification	59
5.4. Accountability and Recourse	60
5.5. Training and awareness	61
5.6. Consumer consultation	62
Appendix 1. The use of AI in finance in Quebec and Canada	63
Asset managers	64
The fintech sector	65
The insurance sector	66
Deposit-taking institutions	67
Appendix 2. Use cases of AI in finance	69
Appendix 3. Lessons learned from deliberation with consumers	77
The consultation process	77
Use case scenarios	78
Summary of consumers’ contributions	83
References	89



Executive summary

Artificial intelligence (AI) technologies are transforming all industries by improving human predictive capabilities, providing decision support or automating certain tasks that usually require natural intelligence. The finance sector is no exception to the deployment of AI systems (AIS), whether in insurance, credit or asset management. The use of AI allows for increased advanced analytics capabilities and thus offers significant competitive advantages that generate interest among financial institutions¹. Financial institutions are encouraged to adapt their operations and organizational structure to take advantage of AI technologies.

While the introduction of AI in finance creates new development opportunities that can be of mutual benefit to financial institutions and consumers, it also creates uncertainties and risks for financial activities, market stability and legitimate consumer interests.

Against this backdrop, the need for oversight of the use of AI in finance becomes pressing. Regulators must exercise their authority to ensure that the use of AI does not distort financial markets, threaten fair competition, or harm consumers.

AI regulation is no longer uncharted territory, but there have been few attempts to effectively regulate AI. The multifaceted and multi-purpose nature of AI hinders management and regulatory projects that must focus on specific applications. It is still difficult to clearly define the issues involved in deploying this set of technologies and applications called “artificial intelligence”.

With previous generations of computing systems, ethical and legal thinking has focused primarily, if not exclusively, on issues of responsible data use and privacy. This approach needs to be adapted in order to face more numerous and varied challenges than those presented by traditional computer systems (non-learning expert systems).

¹ The Authority defines a financial institution as "a company or organization that provides financial services (loans, etc.) to the public and to businesses. Banks, trust companies and credit unions are financial institutions" (see Financial Glossary, https://lautorite.qc.ca/en/general-public/financial-glossary#glossary_F). In this report, however, we refer to "financial institutions" as all companies and organizations that are subject to a framework administered by the Authority. In addition, the recommendations apply to financial institutions in general.

What is AI used for in finance?

In order to better understand the ethical risks that the use of AI creates and to better address the challenges of its responsible deployment, the first step is to have an overview of the uses of AI in finance.

AI, and more specifically machine learning (ML), is generally being used to increase the efficiency and accuracy of operational workflows, enhance performance by supporting multiple aspects of the investment process, and improve the client experience. AI and ML technologies leverage the power of computers and large data sets to derive patterns, to structure non-traditional data sources, and to generate value through effective automation of support and decision-making processes.

A literature review and a series of interviews with financial actors in Quebec and Canada identified four main functions of AI in finance:

1. **Evaluation**, for consumer credit scoring for example, or measuring the ESG footprint of an investment.
2. **Incentive**, to reduce behavioral risks and refine the pricing of insurance products.
3. **Optimization**, for portfolio construction or workflow improvement.
4. **Advice and information**, for personalized customer service, for example.

This taxonomy makes it possible to classify more simply and efficiently the different applications of AI in finance and thus to identify more precisely the ethical issues of the use of AI, which remains very varied.



What are the key challenges to responsible use of AI?

The use of AI in finance raises concerns among consumers who feel particularly vulnerable in a digital and algorithmic environment that they do not always understand and do not have control over. The risks of harm caused by a reckless use of AI in finance must be assessed from two angles: the first one is the degree of harm for an individual, the second one is the magnitude of the impact for a given population. However, a significant harm for an individual can also be inflicted on all the individuals of a population by the effect of large-scale replication of algorithmic decision-making.

Usually, the analysis of ethical issues in the use of AI in the finance sector is based on reports of harm caused by AIS, for example in the area of bank loans, and on the principles of AI ethical charters not specifically related to finance. However, the capacity to report harm depends on the awareness of a harm suffered and the knowledge of the algorithmic processes that are presumably the direct or indirect cause of it. But most of the time, consumers and citizens are not equipped to recognize the harm or its cause.

In order to identify the ethical challenges of AI governance, we proceeded to the consultation of different financial institutions in Quebec and Canada and to the organization of a deliberation process with citizens and consumers, in addition to the state of the art and the analysis of the most significant ethical charters. This deliberation process consisted of a series of workshops to discuss scenarios of AI use cases in finance with which the participants were not familiar. This exercise of collective intelligence helped to go

beyond individual experience and to become aware of risks that might be overlooked in consumers' daily lives.

The interviews with financial institutions and the deliberative workshops with citizens revealed several converging concerns but also somewhat different assessments between stakeholders in the financial sector and consumers. Generally speaking, financial institutions try to anticipate public expectations in order to reduce reputational risks, but they have limited knowledge of public expectations. The value of a public consultation on AI in finance and the process initiated by the Autorité des 9dvent9 financiers ("the Authority") is to improve knowledge of the informed expectations of citizens and consumers.

While financial institutions emphasize privacy and fairness as non-discrimination, citizens are worried about losing their autonomy and their ability to give free and informed consent. The demand for explanation and transparency in algorithmic decision-making aims to restore the conditions for autonomy, whether to consent or to appeal against a decision. For different reasons, this demand from consumers also echoes the concerns of financial institutions, which want to be in control of the decision-making process and prefer, when possible, explainable algorithmic models.

In addition to the more traditional issues of human responsibility and AIS reliability which financial institutions also take into consideration, consumers are notably sensitive to issues of surveillance and privacy infringement, but also to issues of fairness. In this case, fairness refers not

only to the unbiased nature of competitive markets, but more importantly to the non-discrimination of individuals based on irrelevant considerations, and to a kind of social justice. Of course, socioeconomic inequality is not the result of the use of AI in finance, it has broader and

deeper causes, and it is obvious that financial institutions are not mandated to achieve social justice. Nonetheless, consumers caution against the use of AIS that would reinforce socioeconomic inequality, as in the case of consumer credit or health insurance.

Why an ethical framework for AI in finance?

To minimize, mitigate, or eliminate risks associated with the use of AI in finance, a set of principles that promote specific values and goals considered socially desirable is required. An ethical framework establishes the repertoire of shared values and specifies the principles of decision, action and organization. It expresses in a systematic way the expectations of the public, of consumers, and helps to establish a relationship of trust with financial institutions. Finally, an ethical framework supports the adoption of guidelines for AI regulation; it remains valid after the adoption and implementation of guidelines because it allows anticipating what is not yet in the guidelines but could be the subject of new regulations.

Most of the ethical principles of responsible AI can be translated into applicable rules (standards of use, rules of governance), but some of them simply provide guidance. For example, the principle of relevance, according to which one should assess upstream the relevant and

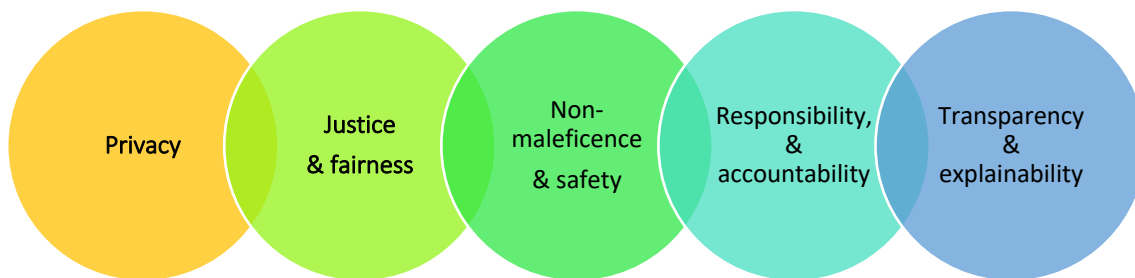
legitimate goals one wants to achieve with AI, provides a direction for financial institutions. Indeed, consumers often seem unconvinced and question the relevance of AI applications. The principle of relevance does not aim to moralize AI in finance, much less to moralize finance using AI, but to underline that any use of a technology by a financial institution should be justified carefully by examining the purposes that the financial institution is pursuing and the means it is using to achieve them.

Most of the ethical principles can be translated and applied through a set of rules, standards, either by degree (*more or less*) or in a binary manner (*obligation or prohibition*). There are now several hundred AI charters and ethical statements and it is necessary to review them in order to establish the principles of responsible AI in finance. But not all have the same quality, robustness or authority.

We have focused on six ethical frameworks in order to identify the normative foundation for developing the most coherent and comprehensive regulation of AI in finance:

- [Montreal Declaration for a Responsible Development of AI](#) (University of Montreal, 2018).
- [Principles to Promote Fairness, Ethics, Accountability and Transparency \(FEAT\) in the Use of AI and Data Analytics](#) (Monetary Authority of Singapore 2018).
- [IOSCO's Fintech Network](#) (IOSCO-IOSCO, 2018)
- [Ethics Guidelines for Trustworthy AI](#) (High-Level Expert Group on AI, European Commission, 2019)
- [Ethically Aligned Design](#) (IEEE 2019).
- [Council Recommendation on Artificial Intelligence](#), OECD/LEGAL/0449(OECD 2019).

By cross-referencing the various documents, a core of five principles can be identified:



However, it is desirable to go beyond the minimal consensus that is established through a cross-referencing of institutional documents to take into account the expectations of the public, citizens and consumers. The *Montreal Declaration for the Responsible Development of Artificial Intelligence* (2018) was developed through a deliberative process that provided a more comprehensive understanding of the ethical issues of AI and proposed a broader set of principles than most other documents. The Declaration takes into account the different values that constitute the minimum consensus above but also imperatives of respecting autonomy, protecting intimacy (not only privacy) and maintaining the conditions of

solidarity, which are crucial for a socially acceptable deployment of AI in finance. For example, autonomy allows us to address the challenges of nudging in the insurance sector; the principle of solidarity also has a specific scope in the insurance sector where the use of AI can lead to hyper-segmentation; the principle of privacy and intimacy allows us to better identify and limit the risks of surveillance.

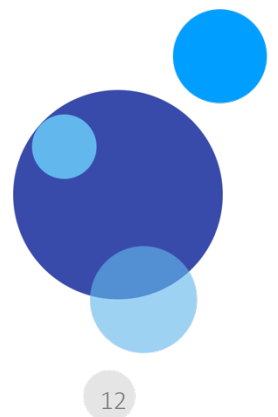
Building on the insights of consumers during the deliberation process organized with the Authority and on the other hand on the *Montreal Declaration* (2018), we have developed a list of principles and standards for the regulation of AI in

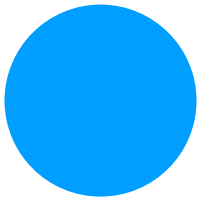
financial markets. This list is a proposal submitted for consideration to the Authority and stakeholders (financial institutions and consumer associations). It leaves many questions open, such as how the regulator could use AI for its

verification operations, for example to check the respect of ESG (Environment, Society and Governance) commitments of financial institutions and to avoid greenwashing that could be detrimental to consumers' interests.

What framework for AI governance?

Ethical principles will not be as effective as intended if financial institutions do not put in place appropriate governance mechanisms. These governance mechanisms include the adoption of a code of ethics for managers and employees of financial institutions, an impact assessment and auditing instrument, certification, accountability mechanisms, recourse and redress procedures for wrong algorithmic decisions, and finally a strategy of public awareness and consultation with stakeholders.





10 Key recommendations

1

The regulator should adopt a model framework for the responsible use of artificial intelligence in finance. By offering values and principles, this framework will allow financial institutions to align their codes of ethics for the use of artificial intelligence and identify unacceptable practices.

2

The regulator should engage in a dialogue with stakeholders to define the sources and types of data that are legitimate to use to operate artificial intelligence systems, regardless of the privacy issue.

3

The regulator must facilitate the training of financial actors in the principles of responsible artificial intelligence in accordance with the normative framework it has established. The regulator must also develop education programs (online resources) for consumers and their associative representatives.

4

Before developing and using artificial intelligence, financial institutions must assess its relevance to their mission and to their customers' expectations.

5

Financial institutions should adopt a governance framework specific to the use of artificial intelligence that allows for human liability for decisions made by an artificial intelligence system, or agreeing with its recommendations, and for accountability for the use of artificial intelligence.

6

To the extent that the use of artificial intelligence significantly increases the volume of decisions and decreases consumer control, financial institutions must adapt their dispute and redress procedures to facilitate consumer action. In the event of disputes, they must offer fast and flexible dispute resolution mechanisms, including mediation.

7

The framework for the responsible use of AI in finance should facilitate the adoption of certification and auditing procedures for artificial intelligence systems.

8

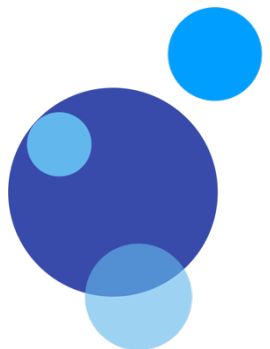
Financial institutions must ensure that artificial intelligence systems meet the resilience principle by being efficient, robust and secure, in order to contribute to the financial market's stability.

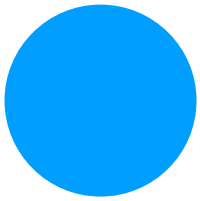
9

Financial institutions must ensure that the use of artificial intelligence systems does not undermine equity, i.e., the equal treatment of consumers, their current or potential customers. In particular, they must avoid reinforcing discrimination and economic inequality.

10

Financial institutions must ensure that the use of artificial intelligence systems respects consumer autonomy by providing all the information required for free and informed consent, by justifying decisions made with the help of algorithms using clear language, and by respecting the diversity of lifestyles.





Introduction

If artificial intelligence (AI) technologies are now developed and applied in all sectors of activity that use cognitive processes of prediction, decision and detection, the finance sector is one of the best suited to their use because of the strong development of financial activities. Whether in the field of insurance, credit or asset management, AI systems (AIS) can automate certain processes or increase the decision-making capacities of human beings. This use of AI to assist human intelligence is sometimes referred to as augmented intelligence.

However, due to the automation of processes in which humans are only marginally involved, and the reliance of human judgment on algorithmic recommendations, the development of AI in finance presents ethical and legal risks for financial

institutions² and raises specific governance issues. Obviously, not all AIS require the same level of oversight as they do not present the same risks to markets and consumers. A distinction must be made between the risks of an AIS that is used to make a decision, such as setting a premium, and those of an AIS that is used to enhance the customer experience, such as automated appointment scheduling.

Many initiatives around the world have been undertaken to establish ethical principles for the responsible use of AI, whether at the level of intergovernmental organizations, governments, industry or academia. Thinking about the framing of AI therefore isn't uncharted territory, but the field of AI regulation largely still is. There are many reasons for this: the proliferation of responsible AI

² The Authority defines a financial institution as “a firm or organization that provides financial services (loans, etc.) to the public or to businesses. Banks, trust companies and credit unions are financial institutions” (see Financial Glossary, <https://lautorite.qc.ca/grand-public/glossaire-financier>). By “financial institutions”, we refer to companies and organizations that are subject to supervision by the AMF, but the recommendations apply to financial institutions in general.

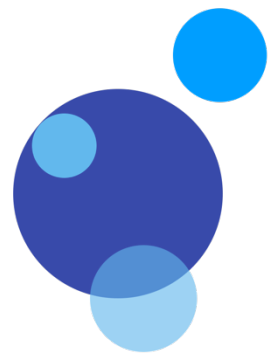
framing documents; competition among stakeholders for recognition of the authority of their ethical principles; the legal or political incompetence of the institution from which the ethical framework emanates (a university is not a regulator, for example); the abstract and limited understanding of ethical principles (it is not enough to evoke an equity or privacy principle for the principle to have meaning); the difficulty of moving from principles to rules and from rules to practices; but also, among other things, the desire to translate principles into rules.

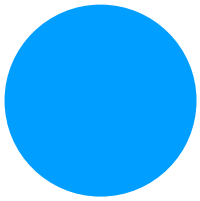
The purpose of this report is to review the use of AI in finance, examine various issues in the responsible use of AI, and make recommendations that would address the challenges of responsible use of AI systems and protect financial consumers. These recommendations are broken down into ethical principles and governance principles; they are not technical standards but may require the implementation of technical standards for the development and use of AIS.

Finally, they must be:

- consistent,
- applicable,
- comprehensive, if not exhaustive,
- acceptable, or even consensual.

Should these recommendations constitute the elements of a binding regulation? This report simply sets out in a coherent manner the proposals for a framework that have emerged from the main international initiatives (intergovernmental organizations and public institutions), from documents published by private financial institutions and from academic publications. In addition, this report was also informed by interviews with stakeholders and, above all, by a public consultation with consumers that was organized in the form of deliberative workshops (Appendix 3). In this way, it differs from other sectoral initiatives in finance. These two consultations with stakeholders and consumers have shown, without theoretical speculation, that private financial institutions favour self-regulation while consumers stress the need for regulation with external verification. They unambiguously expect the regulator to exercise its authority to protect them. However, to the extent that private financial institutions support the implementation of AI ethical principles, their interests and those of consumers are compatible and this report proposes to facilitate their convergence by establishing recommendations that bring together the different stakeholders and satisfy consumer expectations.





2. AI in financial services and investment

2.6. Background and definitions

The last two decades have been marked by the rise of new computer technologies and increased computing power, which, combined with an unprecedented inflation of data sources, have contributed to the development of new applications in a financial sector already open to mathematical modelling and algorithmic risk management techniques. All sectors of activity are concerned, whether it be insurance, asset management, investment, trading, or more globally all functions of a financial institution that involve interaction with a client. Artificial intelligence (AI) and machine learning (ML) are the latest technologies that are attracting the most attention from investors. AI and ML enable the natural evolution of new technologies that leverage computing power and large data sets to identify patterns and representations of knowledge, structure non-traditional data sources, and drive value through efficient scaling and automation of support and decision-making processes.

Today, AI and ML are being used in the financial industry to improve the customer experience, increase the efficiency and accuracy of operational workflows, and improve performance by supporting multiple aspects of the investment process.

In this first section, we provide an initial overview of the uses of AI and ML in financial services and investment. We therefore begin with a brief discussion of the concepts of AI and ML. These two terms are used frequently and very often interchangeably, indicating the existence of many different interpretations of AI and ML. We therefore begin by defining these key terms that give us a basis for discussing AI and ML use cases.³

³ See Nilsson, Nils J. 2010. *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge: Cambridge University Press; OECD/OCDE. 2019. *Artificial Intelligence in Society*. Paris: OECD Publishing., p. 19-35; U.S., National Science and Technology Council, Committee on Technology. 2016. *Preparing for the Future of Artificial Intelligence*. Washington; Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia

Artificial intelligence refers to a field of science that studies and attempts to replicate the various mechanisms that make up human intelligence. This includes various fields such as neuroscience, psychology, behavioural science, biology, anthropology, mathematics, statistics, engineering and computer science. It also encompasses the applied branches of these fields that attempt to replicate human cognition as it manifests itself by performing certain tasks and decision-making processes at performance levels accessible only to humans. These efforts are a concerted combination of computer science and statistical methods that exploit massive data sets and exponentially growing computing power.

Machine learning is a field of computer science that focuses on the problem of designing algorithms and methods that efficiently compress knowledge into a computer system so that it can perform complex tasks via a process similar to “learning” as opposed to hard static programming.

These methods are based on substantial amounts of data from which the system obtains information relevant to the task at hand. Different types of “learning” have been developed for different tasks and analogies can be drawn from the human cognitive process, as some methods can be described as replicating reasoning and learning from experience.

From these definitions, we see that ML is a body of knowledge that allows computers to process data in order to perform certain cognitive tasks. ML is a set of tools and methods by which a computer can “learn” to perform a variety of tasks ranging from simple to complex related to human cognition. In this regard, ML is not a subset of AI but rather a computational approach that studies the question of how to encode knowledge and experience into a computer system through a process analogous to learning. It is therefore a catalyst for AI rather than a subfield.

Hirschberg, et al. 2016. *Artificial Intelligence and Life in 2030.* "One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel." Stanford: Stanford University. <http://ai100.stanford.edu/2016-report>

It is important to raise the importance of the concept of putting an AI/ML model into production. Indeed, in order to generate value from these technologies, these models must be incorporated into a system or work process that links the data to a decision or action in business use contexts. Hence the notion of an AI system.⁴

An AI system is a value-creating process (economic and/or social value, process improvement, task

automation, etc.) that interacts with the context or environment (data and/or human expertise), represents this information through models/methods/approaches, and then generates predictions and/or decisions that, when put into context, serve to increase the efficiency/capabilities of the system's users. We reproduce in Figure 1 a diagram that illustrates the components and connections in a generic AI system.⁵

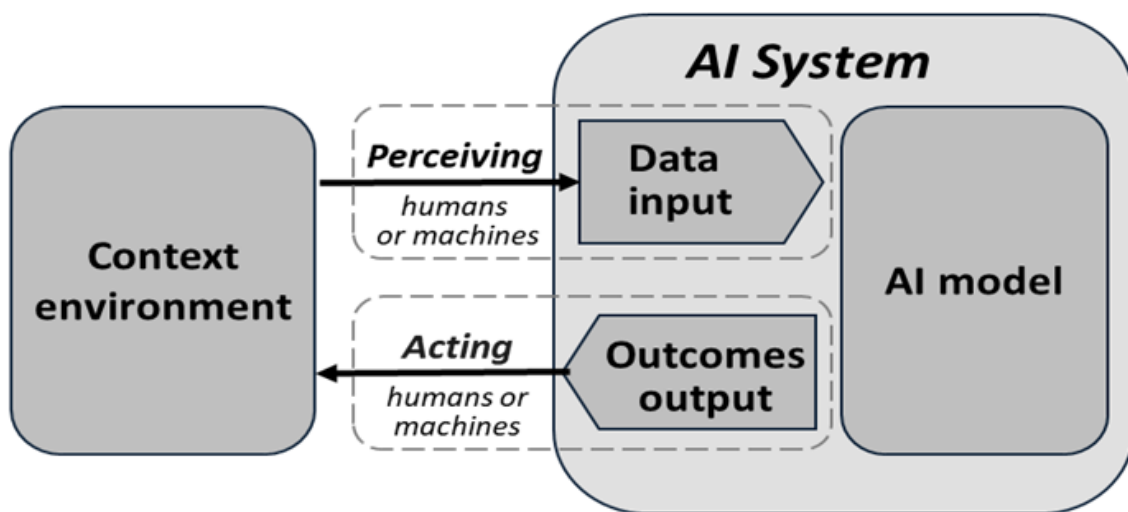


Figure 1: Generic diagram of an AI system. The AI/ML model is only one part in a system whose objective is value creation. Source OECD/OCDE (2019).

The use of technologies commonly classified under the AI label is being used in a variety of ways by financial industry players in their digital transformation initiatives. But two types of applications can generally be recognized: statistical AI and symbolic AI. These notions, which categorize the different models according to the type of data they use as input, also deserve to be

defined in this first section. This distinction will be important in the discussions that follow.

Statistical AI refers to applications that create value or more efficient processes from data through information representation and predictive algorithms that enhance the modelling and predictive capabilities of quantitative teams. These

⁴ OECD/OCDE. 2021. *Recommendation of the Council on Artificial Intelligence*. Paris: OECD/LEGAL/0449.

⁵ OECD/OCDE. 2019. *Artificial Intelligence in Society*, OECD Publishing. Paris: 27-28.

applications automate and/or optimize existing investment decision-making processes.

Symbolic AI refers to applications that create value from the representation of human knowledge and expertise; this acquisition process leverages unstructured information, which is then transformed into structured and predictive data to increase the opportunities for value creation from the data. These applications create new opportunities for automating and/or optimizing

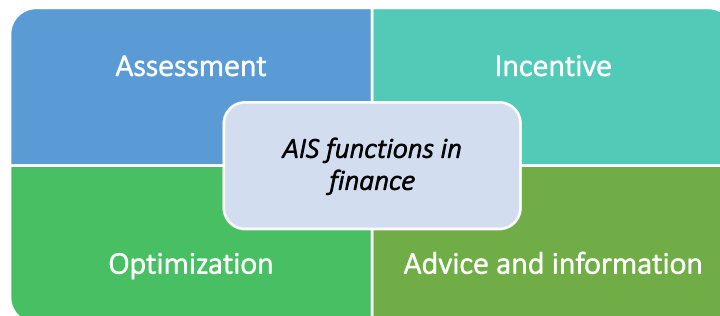
decision-making processes in places not yet explored by statistical AI.

Having defined the concepts and notions fundamental to the use of AI and ML for value creation, we present in the following section a first categorization exercise of these applications. We propose a typology by function that will be useful when analyzing the risks and impacts of using these models in the financial services and investment sector.

1.2. Using AI in finance and investment: taxonomy by function

The last few years have seen a significant increase in the adoption of AI/ML technologies in financial services and asset management. There are a growing number of use cases, proofs of concept, solutions and even products that use these technologies to varying degrees. These AI systems are used with different objectives and AI/ML methods along the value chain of institutions operating in the financial services and asset management sector. It will be important for the discussion to create a taxonomy that classifies the different use cases or applications of these technologies. This taxonomy should be the basis for the subsequent analysis that will address the ethical risks and impacts of production deployments of these AI systems. Existing taxonomies focus either on the type of underlying models or techniques, or on the types of data used, or on the business context where the system is deployed. We propose here a taxonomy that would be based instead on the function that an AI/ML technology element performs within an AI system. In this taxonomy, a use case can be seen as the use of an AI system in a business context, which in turn will be composed of several functions that call upon specific technological elements.

We propose a taxonomy that breaks down four functions of AIS in finance:



We present, in boxes, a use case for each function⁶. These boxes illustrate how a business problem is addressed in an innovative way by deploying an AI system that generates value for the company and users through digital strategies, data and AI/ML models. An AI system usually performs more than one function at the same time.

1.2.1. Assessment

A technological element of an AI system performs an evaluation function when it leverages an AI/ML model or technology to generate a quantitative measure from the data. These measurements are then used at another level of the AI system including its final output. The evaluation function thus allows to represent a complex reality in a finite set of measurements, effectively enriched data, which will feed other steps of an AI system. Examples of elements of an AI system that performs an evaluation function can be found in several use cases.

Here are some examples:

Credit scoring. In an AI system that has credit adjudication as its objective, there are models that transform an individual's historical characteristics (transactions, buying preferences, risk tolerance, debt level, etc.) into a score that is used further in the system to determine the risk associated with that individual and potentially make a decision on granting credit and its parameters. Scoring systems are also used in the insurance industry for underwriting purposes.

Fraud detection. In an AI system that aims to detect fraudulent transactions in an account (credit or current), one can find models that, using a history of transactions, personal data and metadata⁷ collected during user interactions, will generate a qualitative measure of the level of veracity or authenticity of transactions. This measure is used later in the system to make a decision on the authorization of a transaction.

ESG Footprint. In an AI system that aims to integrate environmental, social, and governance (ESG) factors into an asset manager's investment process that aligns with sustainable development goals (SDG), one can find models that will quantitatively represent the information contained in companies' ESG disclosure reports as well as the original data sources of these reports to inform the decision process. This qualitative data is unstructured and AI/ML models can transform it into metrics that allow it to be integrated into fundamentally quantitative investment processes.

Pricing of auto insurance products. In an AI system that has the objective of calculating an auto insurance premium, there are models that use an individual's personal data, driving history and real-time road behaviour, to generate a quantitative measure of their associated risk. This measure can be used further in the AI system to build a final rating.

⁶ See Appendix 1 for other documented use cases.

⁷ By metadata, we mean any data that describes another data, characterizes it or provides information about its nature in order to facilitate its understanding, classification and management.

Case 1. *Lenddo* and the credit score

Field: Credit

Primary use: Assessment

Secondary Use: Advice and information

Context

Traditionally, building credit to achieve a favourable credit score takes time and financial activity. On the creditor side, credit scoring is also time consuming, as a lot of personal data must be taken into account, such as income, payment history and possibly the credit history from other financial institutions. In order to offer an alternative to traditional credit scoring, *Lenddo* has created an application that uses transaction data as well as non-financial data to give access to credit to customers with no credit history.

Description and objectives of the system

Lenddo overcomes the slowness and labour of the traditional credit system by offering online and mobile services to calculate a credit score for users with no credit history. They can simply install the *Lenddo* application on their smartphone, and *Lenddo* will determine their credit score by analyzing their digital footprint.

This digital footprint is based on financial transaction data, activity on social networks and search engines, and geolocation data. *Lenddo* also makes judgments based on information it infers from analysis of its users' behaviour. For example, it considers things like the tendency to write more than one word in the subject line of an email (which would mean that customers are detail-oriented), and regularly using financial apps on their smartphones (which would show that the user takes their finances seriously). *Lenddo* is also looking at the ratio of photos in the photo library taken with a front-facing camera, as self-portraits would be an indicator of youth, which would help the company segment its customer base.

System benefits

Banks and credit unions can then use users' *LenddoScores* to better understand their risk of defaulting on loans. Thus, with access to a larger volume of customer data, banks with this technology can provide credit to people who would not otherwise have access to it.

Technology used

The software uses natural language processing to analyze users' social media posts and what they type into their browsers for indicators of liability or risk-taking. Then, this information informs predictive analytics algorithms that create a credit score. Predictive analytics assess whether or not a customer is likely to repay.

1.2.2. Nudge

An AIS performs an incentive function when it leverages an AI/ML model or technology to generate interactions with the system's end user that seek to influence their behaviours and give them a particular direction. This influence seeks to facilitate the adoption of the system, encourage its use or generate other sources of data that can be used by the system to improve its performance and/or for the value generated for the business context in which the AI system is deployed. Examples of the elements of an AI system that perform an incentive function can be found in several use cases, whether in the insurance or credit field.

Here are some examples :

Pricing of car insurance products. In an AI system whose objective is the calculation of a car insurance premium, one can find AI/ML processes whose objective is to influence the users of the system, via personalized rewards or “nudges”, towards a safer behaviour on the roads. This function helps personalize the price but also minimize the overall losses.

Solution to assist in the management of current accounts. In an AI system whose objective is to assist in the management of bank accounts (current or savings), we can find AI/ML processes whose objective is to recommend savings options to customers, to signal when they exceed their budget, or to suggest simple methods to reach their financial goals. This feature helps personalize clients' expectations of their budgets and savings goals, encourage new methods, suggest new products, and generally create a healthy and appropriate financial literacy. By providing an overall picture of an individual's personal finances, the solution seeks to encourage saving and investing through other products.



Case 2. *Vitality* by John Hancock: the gamification of life insurance

Field: Insurance (Health)

Primary use: Incentive

Secondary use: Advice and information

Smart devices (SD) ⁸ offer great potential to insurers who want to take precise measurements of the parameters they monitor when calculating their customers' insurance premiums.

John Hancock Life Insurance Company (JH) – a subsidiary of Canadian insurance company Manulife Financial and operating in the life insurance, financial advisory and wealth management business – created their first SD-based interactive life insurance policy in 2015: the *Vitality* program, created with their exclusive partner Vitality Group. Starting in 2018, they announced that they would now only sell interactive life insurance plans. With the SDs included in their plans, JH is able to base their pricing on the data they capture on the devices their customers carry, as well as their smartphones. Holders of these insurance policies are able to get discounts on their premium if they meet exercise targets that their SDs record. They also get special prizes such as gift cards from various stores if they register their workouts and healthy food purchases in the *Vitality* app.

Description and objectives of the system

The logic behind such a precise calculation is one of efficiency. What Brooks Tingle, CEO of JH, describes as a “virtuous cycle” is explained in these words: “If our customers take steps to live longer and healthier lives, it definitely creates value for us, and we don't hesitate to recognize that. People live longer, we make more money. In short, customers are encouraged to take care of themselves by staying fit. They get the natural benefits of being fit, plus the monetary and merit benefits from JH, who can make money longer with the same client while ensuring that their actuarial calculations are most accurate for each client.

System benefits

The value created by this business model is such JH can afford to give back to their customers in various ways to acknowledge their participation: getting a free SD, discounts on premiums, up to 25% off their life insurance premium, significant discounts with certain retailers. Customers can also receive 25% off their purchases. The most engaged customers can even earn a free subscription to services like Amazon Prime.

JH uses a *gamification* strategy to encourage its clients to adopt healthy behaviours and habits. Gamification is a motivational strategy that uses game mechanics to encourage a person to perform certain actions. More

⁸Also known as the “*Internet of Things*” (*IoT*). Smart devices are an ecosystem of devices capable of joining a network to receive and send data to and from objects via the Internet.

specifically, JH uses a gamification strategy that seeks to foster a sense of accomplishment in its users by giving them rewards.

JH customers can choose between a plan that allows them to log their activities in an app or website and receive gift cards, or take a more ambitious plan that entitles them to a wearable SD, such as a smart watch (*FitBit*, *Apple Watch* or *Amazon Halo*), as well as discounts on rewards and other types of rewards. Customers also have the opportunity to play a wheel-of-fortune type game when they reach certain goals. Again according to Brook Tingle, this is one of their customers' favourite perks: "[...] you go to your mobile device and there's a wheel that looks like a Wheel of Fortune – or something like that – and you spin it [...] you can earn \$5 at Starbucks, \$10 at Amazon, and the most consistent positive feedback we get isn't, "Oh, I saved \$1,000 on my premium, that's great!" [...] it's, "I love this wheel!"

Technology used

SDs primarily use machine learning (ML) to analyze data points from sensors, accelerometers, gyroscopes and GPS from an API⁹. Commonly used ML methods are: decision tree forest algorithms¹⁰, J48 decision tree algorithms, learning from examples¹¹, Bayesian I classification processes, and multi-layer perception algorithms.¹²

⁹ *Application Programming Interface*. "The API can be summarized as a computing solution that allows applications to communicate with each other and exchange services or data with each other." (4)

¹⁰ *Random Forest* (RF)

¹¹ *IB3 Instance-based learning*

¹² *Multilayer Perception* (MLP)

1.2.3. Optimization

A technology component of an AI system performs optimization when it leverages an AI/ML model or technology to find an optimal solution to a specific problem or process under defined constraints. This use seeks to facilitate realistic results, use more suitable methods, and incorporate other data sources that can be used by the system to improve its performance. This function also includes AI systems that optimize a process by automating procedures in a workflow that can range from data collection to decision-making. Examples of the elements of an AI system that performs optimization can be found in several use cases.

Here are some examples:

Portfolio building. In an AI system whose objective is portfolio building, we can find AI/ML processes whose objective is to find an optimal allocation of the securities included in a portfolio while respecting certain constraints of return objectives and risk tolerance. This function helps identify new opportunities in the bond and equity markets, adequately diversify the portfolio by identifying new correlations, or reduce model calibration and parameterization errors by using model-free algorithms (e.g., reinforcement learning) in order to more precisely identify an optimal portfolio that meets the needs. An AI system can facilitate the integration of new factors (macroeconomic conditions, economic cycles, etc.) in the analysis and prediction as well as unstructured data such as textual information in financial reports, social media, blogs or news through language processing and sentiment analysis.

ESG Footprint. In an AI system that aims to reduce the ESG footprint, one can find AI/ML processes whose objective is to find an optimal allocation of securities included in a portfolio with a primary focus on reducing the ESG footprint. This function can identify new opportunities in the bond and equity markets, and validate the information disclosed by certain investments through the use of unstructured data such as textual information in financial reports, social media, blogs or news through language processing and sentiment analysis.

Workflow optimization. In an AI system whose objective is workflow optimization, one can find AI/ML processes whose objective is to optimize the “path” to perform a given job. This can be used to minimize costs, automate processes, or maximize output and employee productivity. This could apply to issues of logistics, supply chain, cash flow, routing of interbank transfers and payments, or even optimal broker selection.

Case 3. V.I.T.A.L. on the board of directors of Deep Knowledge Venture

Area: Investment

Primary use: Optimization

Secondary use: Advice and information

Context

VITAL (Validating Investment Tool for Advancing Life-Sciences) was an algorithm named to the board of Hong Kong-based venture capital firm *Deep Knowledge Venture* (DKV) in May 2014. DKV specializes in regenerative medicine, and VITAL was specifically helping to evaluate potential investments in biotech start-ups. It had a vote on the company's investment decisions in the same way as a human member. It hasn't served since 2017.

Description and objectives of the system

VITAL was developed for DKV by Aging Analytics UK. The purpose of the program was to predict, for DKV's board members, which companies and technologies in development were most likely to offer a profitable return on investment. Since nine out of ten biotech start-ups fail, VITAL has proven to be a useful tool for detecting the warning signs that a company is in a poor position and is too risky an investment.

Benefits of the model

VITAL's analytics gave DKV a strategic investment advantage to position itself as a leader in biotechnology and regenerative medicine. Using AI allowed DKV to make rational, fact-based business decisions and to move away from the emotions that can colour a human's judgment. However, DKV's CEO also recognized the benefits of intuition, so VITAL's logic was used to support this human quality, not replace it.

Technology used

VITAL evaluated fifty parameters, such as share price, clinical trials, ownership of intellectual property rights, or research funding obtained, but little information is available on its programming.

1.2.4. Advice and information

A technology component of an AI system provides advice and information when it leverages an AI/ML model or technology to advise or create value to assist in decision-making. This function seeks to bring out otherwise hidden information, in order to clarify, enhance, or improve existing information, in order to provide better service to consumers. Examples of the elements of an AI system that performs an advisory function can be found in several use cases.

Here are some examples:

Personalized customer service. In an AI system that aims at personalizing customer services, one can find AI/ML processes whose objective is to better understand customers and adapt its advice and proposals for new offers. This function helps identify the hidden needs of customers and create new business opportunities for financial institutions. These personalized, anywhere, anytime financial services are increasingly possible thanks to mobile applications that are changing the way customers interact with financial institutions, and the analysis of data from the financial institution's customer smart devices (point-of-sale devices, GPS tracking, inventory tracking). This understanding of the customer also enables customer insight to more easily detect unusual behaviour and fraud.

Investor advice. In an AI system that aims to support investors' choices internally, one can find AI/ML processes whose objective is to offer investment recommendations based on new and more easily exploitable data sources (financial reports, social media, news, blogs), sentiment analysis towards securities and their trends, and in-depth analysis of portfolio managers' behaviour. This feature helps advise and build customized dashboards for portfolio managers.

ESG Footprint Reporting. In an AI system that aims to advise companies in their ESG reporting, one can find AI/ML processes that aim to find a way to disclose ESG information in a way that highlights the steps taken towards ESG aspects. This function ultimately changes the perspective of investors, and also the public, towards some companies that otherwise have difficulty publicizing their efforts. This leads to a transparency of the measures taken in this context.

Case 4. *Truvalue Labs* and ESG disclosure

Area: Investment

Primary use: Advice and information

Secondary use: Optimization

Context

ESG factors are criteria used to measure the sustainability and societal impact of a company or investment. The acronym ESG refers to environmental, social, and governance issues. From an environmental point of view, concerns are about the impacts of economic activities on global warming and their environmental sustainability. On the social level, issues of diversity, equity and respect for human rights and the protection of people are at the forefront. And finally, at the corporate governance level, there is an interest in responsible management structures, employee rights, management and responsible use of private data. *Truvalue Labs*, an analytics and investment support company, use AI to discover opportunities and risks hidden in massive volumes of unstructured data to analyze, track and score companies' ESG compliance in real time.

Description and objectives of the system

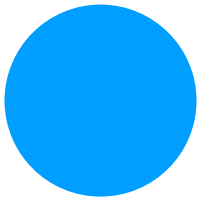
Truvalue Labs seek to identify common metrics and consistently report on sustainable value creation from a company's ESG performance indicators. *Truvalue Labs* identify material issues, quantifies them, and then produces a suite of data and analysis including scores, trend graphs and sources. This information will help investment firms analyze and interpret massive amounts of unstructured data at high speed, enabling them to make informed investment decisions.

System benefits

This use of AI allows for millions of data points per month. It has the potential to replace multiple analysts, while being more efficient, given the amount of data that can be quickly analyzed. In an extensive back test, *Truvalue Labs* proved that incorporating data that focuses on a company's ESG criteria can generate above-average returns.

Technology used

This platform uses two different AI technologies to achieve its goals: natural language processing and massive data analysis. Thus, they are able to capture unstructured data from thousands of sources on the Internet, 24 hours a day, and generate high-value insights for investment firms.



2. Ethical risks and challenges of responsible AI deployment in finance

The use of AI in finance offers mutually beneficial opportunities for both financial institutions and consumers. However, it is now well established that this use presents ethical and legal risks for financial institutions and raises concerns among consumers who feel particularly vulnerable to algorithmic decisions they cannot control. These concerns are not simply knee-jerk reactions, but judgments informed by ethical values and legal considerations for rights.

The consultation of various financial institutions in Quebec and Canada and the organization of a deliberation process with consumers (see appendices 1 and 3 of this report) have made it possible to identify the risks that financial institutions are trying to anticipate and the ethical and societal concerns of consumers. We present the views of financial institutions and consumers in turn, not to oppose them but because, due to their situation, they have different AI experiences.



2.1. The perspective of financial institutions

Financial institutions generally have competent staff responsible for the development or deployment of information technologies, data management and use, and for the largest of them, a legal department, with ethical qualifications, specialized in information technologies. That is usually accompanied by the presence of a strong multidisciplinary team and a desire to combine expertise in order to understand IT development, especially AI in its entirety: computer scientists, mathematicians, lawyers, ethicists, managers, etc. are called upon to break down silos, work together and develop a common language. It is therefore not surprising to see IT development, modelling or Innovation managers able to answer questions on the responsible use of AI And legal department managers talking about their company's latest AI tools.

However, while awareness of ethical and legal issues is acute within these financial institutions, it is noticeable that it focuses on a limited number of ethical and legal issues. Two issues receive the most attention :

- Protection of personal data ;
- Equity as non-discrimination.

2.1.1. Privacy

Privacy law is the primary focus of financial institutions and the reason for this is twofold: first, it is a legally recognized and enforced right, and therefore it is imperative to respect it regardless of one's moral motivation to do so; second, it is a law with a long history that is independent of recent AI developments and is therefore more familiar.



2.1.2. Equity

Equity is a concept that will be discussed in more detail in section 4.3. It is generally understood as the value that prohibits unjustified discrimination, and the latter notion is reduced to the equal treatment of people according to their community affiliation defined by race, religion or gender. The social justice dimension is most often (with rare exceptions) neglected in favour of a formalistic interpretation: the use of AI is fair if its consequences are neutral, i.e., if it does not

disproportionately affect members of a given community.

We cannot exclude that the reasons for this interest in these two ethical-legal issues are moral, in the sense that financial institutions (the ones responsible for deploying AI) take consumers' ethical expectations seriously, but it clearly appears that they wish to avoid legal and reputational risks that would strongly penalize their activities.

2.2. Consumer expectations

Consumers also identify the two issues of privacy and fairness as non-discrimination, but their perspective differs in that it asserts an ethical interpretation of their interests and rights. Above all, they have an interest in a diversity of values and principles that are not always taken into consideration by financial institutions. If the main issue for financial institutions is to avoid legal and reputational risks, the main issue for consumers is the conditions of trust in AI, that is to say, in the use of AI by financial institutions.

How can we trust AI in finance? Consumers generally understand the benefits of AI to improve their customer experience, protect them from fraud, and optimize their investments. They also understand the interest of financial institutions to reduce their costs and increase their margins by automating some processes that used to be done only by humans. But they have ethical interests as individuals and as citizens and have expectations of financial institutions to ensure that these interests are respected.

Their expectations focus primarily on four issues: respect for autonomy, promotion of their well-being which includes respect for their privacy, commitment to social equity and human responsibility¹³.

¹³ See Appendix 3 for a more detailed description of consumer expectations as they were formulated during the deliberative process organized by the Authority with the University of Montreal.



2.2.1. Respect for autonomy

First, consumers are concerned about a loss of autonomy in the face of AI deployment at financial institutions. They believe that the lack of explanation and transparency in algorithmic decision-making, or human decision-making augmented by the use of AI, may constitute an infringement on their ability to consent in a free and informed manner. The intelligibility of algorithmic decision and processes is a key requirement for consumers. Consumers also expect financial institutions to be able to justify decisions that affect them in non-technical language that is accessible to non-specialists. This implies that managers in financial institutions are able to understand and explain these processes. The explainability of AIS, their intelligibility and transparency are not independent ethical values; rather, they correspond to properties of AIS and their institutional environment that make it possible to respect consumers' autonomy.

Consumers are particularly sensitive to the issue of autonomy in relation to assessment AIS, for example, for credit and for nudging in the health insurance field. Nudging techniques are sometimes seen as paternalistic methods of coercing customers to do what is in the interest of the financial institution and not what is in their own interest. They expect financial institutions to strike an appropriate balance when inducing and counselling between the financial interests of the institutions, on the one hand, and their moral interest in exercising judgment and maintaining their lifestyle on the other. It is not enough for financial institutions to demonstrate that their financial interests coincide with the financial interests of consumers, because even if the latter are sensitive to this argument, they wish to maintain the conditions for exercising their autonomy. The consumer is above all a person with moral interests.

2.2.2. The risk of surveillance

We will return to the practice of nudging in detail (Section 4.5) from a normative perspective, but it is important to emphasize here the surveillance issue that pervades thinking about consumer AI. Consumers are concerned about the deployment of intrusive AIS either because they are powered by intimate personal data or because they act as moral guardians. This is the case with smart devices used to assess policyholders' lifestyle habits. In addition, some AIS developed in the credit sector use unconventional data such as Internet browsing history, messages shared on social networks, location data or even images (selfies).

This type of AIS creates a feeling of being watched and therefore affects the well-being of consumers but also their autonomy by exerting pressure on their lifestyle and judgment. This is why consumers are very sensitive to the protection of privacy (and their intimacy). While transparency of AIS and their institutional environment is important to ensure the proper use and governance of AI tools, it should not reach the privacy of consumers who wish to protect their sphere from the intrusion of some kind of algorithmic gaze.

2.2.3. Equity and solidarity

While the potential for algorithmic bias and discrimination on the basis of race, religion, or gender is well identified as an infringement of equity, consumers pay attention to another dimension of equity: socioeconomic inequality. Bias and discrimination on the basis of race, religion, or gender obviously have socioeconomic consequences, but they are often treated as a formal equality violation. Yet even if they were neutralized, socioeconomic inequalities that have more global causes would not be reduced within each group and could be amplified by a range of social determinants such as access to digital infrastructure (e.g., cell phones), by education and digital literacy, by health, or by behaviours considered by the majority to be risky.

While socioeconomic inequality is not a product of the use of AIS in finance and financial institutions are not in the business of achieving social justice, consumers point to uses of AI that unduly reinforce socioeconomic inequality: for example, assessing policyholders' health through connected sensors coupled with incentive AIS to adopt healthy behaviours can exacerbate the effect of social determinants of health and disadvantage policyholders who are not able to follow the app's recommendations.

In the credit arena, consumers are concerned about another practice of using unconventional data to assess the solvency of people who do not have a credit history or who do not qualify for more traditional credit programs. Smart apps embedded on phones are thus being deployed by financial institutions and other third parties (fintech) to enable young customers in particular to borrow. While this may appear to be advantageous to would-be borrowers, the risks of going into severe debt for individuals who are vulnerable to these online offerings and ill-

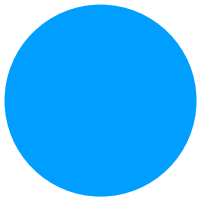
equipped to understand the solutions offered by AIS are high.

2.2.4. Responsibility

Finally, consumers expect that the deployment of AIS in financial institutions and decision-making should always be a human responsibility. There are at least two ethical reasons for this. The most obvious is that AIS can hardly be considered

morally responsible for the decisions made. Although automated, intelligent machines cannot be said to be autonomous. Only human beings and legal entities (institutions) can be held accountable for decisions made whether or not they come from AIS. The second reason is that consumers have a moral interest in maintaining a human connection with their financial institution and being able to identify the people within the institution who will justify the decision affecting them and be accountable for errors in judgment and management.





3. Assessing the goals of AI in finance upfront

When the public is presented with AI use cases in finance, a common reaction is to ask why a particular AIS was developed or what type of data was used. Consumers often appear incredulous and question the relevance of these AI applications. Before examining the ethical principles of responsible AI, it seems appropriate to assess the goals that the use of AI in finance should achieve. This is not to moralize AI in finance, let alone to moralize finance using AI, but to emphasize that any use of a technology by a financial institution should be preceded by a reflection on the ends that the financial institution pursues and the means it implements to achieve them. This is especially important as the risks of misuse or adverse consequences for consumers increase. In the case of AI applications, the level of risk to individuals and the magnitude of adverse consequences for entire segments of the population can be considerable. A general argument supports this approach: financial institutions are particularly sensitive to reputational risks in the event of misuse of AI or AI that is rejected by consumers. It is important to ensure that the goals of AIS deployment in finance are acceptable to the public.

3.1. The relevance principle

If the technology presents significant risks, its use must be *justifiable* to consumers or regulators. To put it another way, the objectives must be relevant to the financial institution's business, and the means, the AIS, must be relevant to achieving those objectives. The use of an AIS or data set is relevant if it enables the financial institution to perform a function, such as reducing investment or insurance risks.

In order to establish the appropriateness of deploying an AIS, it is imperative that the mission of the financial institution be very clearly defined. Doing so avoids moralism and allows the relevance principle to be applied with caution.

Let's take two examples:

Ex.1. As part of its mission, an insurance company must be able to respond to customer requests for information. If the company deploys an AIS as a conversational agent whose purpose is to provide information to customers, the deployment of this AIS is relevant. Other constraints will need to be considered later, but from a relevancy perspective, the AIS qualifies.

Ex. 2. As part of its mission, a property and casualty insurance company must be able to predict the risks faced by customers in order to be able to compensate them in the event of a loss. If the company deploys an AIS that aims to change the lifestyles of customers in order to reduce the risk of harm, the objective does not seem to be part of the financial institution's mission. But as a risk specialist, the insurer can guide the consumer and advise them on how to control the risks to which they are exposed.

The application of the relevance principle can be broken down into four steps:

- a. Check that the AIS **achieves a financial institution's mission objective**;
- b. Check that its deployment **promotes the interests of the parties** (the financial institution and its clients);
- c. Check that the **benefits of AIS deployment outweigh the drawbacks**;
- d. **Assess the type of benefit** it provides for each party and check that it is more desirable than the type of risk or inconvenience.

3.2. The resilience principle

In finance, AIS can only be relevant if they don't undermine the stability of financial markets, or if they reinforce it. We must therefore ensure the resilience of the AIS used, a resilience principle that we define in three criteria: efficiency, robustness and security.

Effectiveness means that, depending on the preferred metric (profit, compensation, client information, dispute processing time, etc.), the use of AI leads to better outcomes than without its use. This is difficult to prove, but it is possible to satisfy this kind of counterfactual criterion by implementing an evaluation protocol with control samples.

Effective AIS must also be robust. The criterion of robustness refers to the idea that AIS should behave predictably under different circumstances, both favourable and unfavourable, and not cause unanticipated harm and damage. In other words, one must ensure that the operation of AIS achieves the objectives that have been set and only those objectives. This also rules out the possibility of double use of the machine by achieving the programmed objective (e.g., predicting risks) at the same time as collateral objectives (e.g., monitoring customer behaviour, micro-targeting advertising).

Finally, it must be ensured that the AIS deployed are secure and cannot be deceived, especially by sabotage or data poisoning techniques used to train or operate the algorithms. If the quality of the data sets is degraded, this leads to unpredictable and unstable AIS that are no longer able to function according to their programmed goals, rendering them useless at best and malicious at worst. Given the volume of decisions that AIS can make in finance, this unpredictability could affect the efficient functioning of the financial markets themselves and even increase systemic risks.





3.3. AI and the development of responsible and sustainable finance (ESG)

To ensure long-term value creation and stability in potentially turbulent markets (just think of environmental or health risks), financial market participants, including investors, are increasingly adopting ESG (environment, society and governance) criteria to guide their activities and select financial products. Major companies, under the aegis of the World Economic Forum’s International Business Council, have joined forces to catalyze the widespread adoption of common ESG metrics, aligned with the United Nations’ 2030 Sustainable Development Goals. Moreover, while the goals of sustainable finance meet the ethical

and social expectations of consumers and investors, it should be noted that they also increase financial performance and support equity in financial markets by framing the competition between financial institutions focused on quarterly balance sheets and tempted by short-term value creation¹⁴.

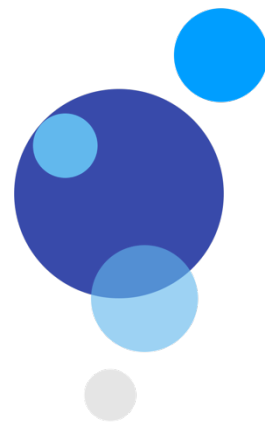
Against this backdrop, financial institutions are increasingly recognizing the transformational role that AI technologies will be called upon to play in the coming years to help them implement ESG standards in their financial activities. AI investment

¹⁴ We note, however, that *greenwashing* and *virtue signalling* are omnipresent and blur the signals and harm the efficiency of this market. The development of ESG labels and standardized ESG indices will contribute to improving the quality of information made public.

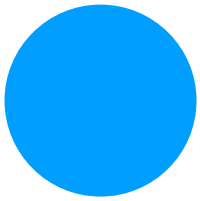
applications incorporating sustainable finance goals is one of the most important innovations in finance and should be combined with applications to automate and optimize decision-making processes to provide desirable guidance to various financial market participants.

Noting this trend of financial institutions engaging with their clients to promote “green” and socially responsible investments¹⁵, the regulator should be

able to ensure that these ESG commitments are respected and do not constitute “greenwashing” as it affects the decisions of their partners and clients. Just as AI can help financial institutions comply with ESG standards by assessing and tracking investments, it can be used by the regulator to verify the ESG commitments of these institutions, for example by using automated reporting techniques that analyze large amounts of data.



¹⁵ See, for example, Blackrock’s decision in favour of a zero-carbon transition: “BlackRock pushes companies to adopt 2050 net zero emissions goal,” Michael Mackenzie and Billy Nauman, *Financial Times*, New York, 26 January 2021. <https://www.ft.com/content/a71feaac-d3f4-4e76-a60c-c68924b06dfd>



4. Ethical compliance principles

After determining the objectives of a relevant use of AI in finance, one must ensure the ethical compliance of the AIS and its use by first identifying the ethical principles that apply to the development and deployment of AI. These ethical principles must be applied at the time of the design of the AIS by the IT developers (ethics by design) and in the context of its deployment (ethics in use).

But why import the discourse of ethics into the language of business? Why not be satisfied with legal compliance? Let's remember that the law is not always adapted to the evolution of new technologies and in particular in the case of AI where it is still very timid and evolves according to new needs to sanction or prevent new types of harm. However, because of its inadequacy and its evolution, it cannot prevent all reputational risks for financial market players. Thus, without ethical guidelines, the law's domain of certainty does not resolve the domain of uncertainty and instability in financial activities. A guideline may provide more leeway to manage AI than a law or regulation.

Above all, it is important to remember that the law is the law, and that legal compliance is a legal, not an ethical, obligation imposed on all financial

market participants. Depending on the sector of activity, this applies to provincial and federal laws, but also to international legal provisions such as human rights conventions. There is no need to stress that financial institutions, in developing, adopting and using AIS for their activities, must scrupulously observe the applicable national and international legal provisions. When we speak of human rights as they are formulated and implemented in positive law (conventions,

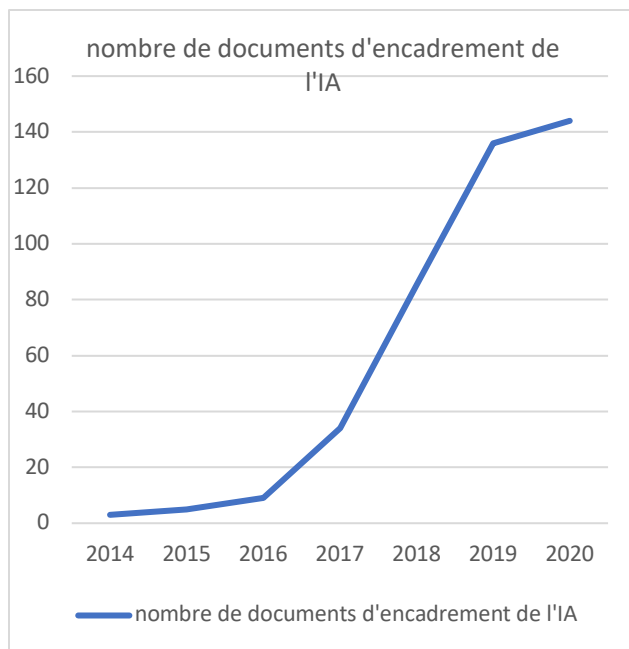
international treaties), we are referring to positive legal obligations¹⁶.

The challenges for the deployment of relevant, responsible and sustainable AI in finance go beyond the framework of law, national or international, and it is not enough to apply the law as it exists to address them satisfactorily. Consumer protection requires more than the non-violation of human rights under international law and compliance with existing national laws.

Financial institutions can only reduce reputational risks and prevent harm from misuse of AI by anticipating them through guidelines or ethical principles. These principles express in a systematic

way the expectations of the public, of consumers, and help build trust in financial institutions.

Note that the quest for efficiency by financial institutions is generally guided by principles that go beyond the legal framework and that, in order to be able to identify risks of misuse of AI, as they do, it is necessary to have an ethical evaluation grid, even if this grid is not formalized as such. In this report, we try to propose a formalized framework. By going back and forth between intuitively identified risks, consumer expectations and formalized ethical principles, it is possible to refine the ethical assessment and better cover all the risks of AI use by financial institutions.



In a few years, the number of proposals in artificial intelligence ethics (declarations, guidelines, charters, recommendations, etc.) has increased considerably. Based on data from the algorithmwatch.org directory site¹⁷, we count **144 documents in 2020** whose evolution we represent in the table below (we estimate that the real number is higher, because several references are missing from this list, such as the Declaration of the Global Partnership on AI, but also many documents from the private sector).

¹⁶ When human rights are mentioned simply as ideals to be respected, they consist of guiding ethical principles, principles that underlie the evolution of law.

¹⁷ "AI Ethics Guidelines Global Inventory," accessed May 18, 2021 at <https://algorithmwatch.org/en/ai-ethics-guidelines-global-inventory/>

It is not easy to find one's way around. However, not all of these proposals have the same ethical value, function or target audience. While corporate statements¹⁸ are very useful for identifying the concerns of private sector financial institutions, they are too numerous and do not have the same scope or independence as documents from international organizations, financial authorities and the academic sector. For the private sector, we will present the principles that we were able to gather from our survey of a variety of financial institutions in Quebec and Canada.

Outside of the private sector, several ethical frameworks of reference deserve particular attention. Among these frameworks, the *Montreal Declaration for the Responsible Development of AI*¹⁹ will be favoured because of its more comprehensive nature and its deliberative elaboration open to public participation.

4.1. Frameworks for the framing of AI in the financial sector

In chronological order, our preferred frames of reference are as follows:

- i. [Montreal Declaration for a Responsible Development of AI](#) (University of Montreal, 2018).
- ii. [Principles to Promote Fairness, Ethics, Accountability and Transparency \(FEAT\) in the Use of AI and Data Analytics](#) (Monetary Authority of Singapore 2018).
- iii. [IOSCO's Fintech Network](#) (IOSCO-IOSCO, 2018)²⁰
- iv. [Ethics Guidelines for Trustworthy AI](#) (High-Level Expert Group on AI, European Commission, 2019)
- v. [Ethically Aligned Design](#) (IEEE 2019).
- vi. [Council Recommendation on Artificial Intelligence](#), OECD/LEGAL/0449(OECD 2019).

¹⁸ For instance, IBM. 2019. *Everyday Ethics for Artificial Intelligence*. IBM. <https://www.ibm.com/design/ai/ethics/everyday-ethics/>; Microsoft. 2019. *Responsible AI Principles*. Microsoft. <https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimar6>

¹⁹ The *Montreal Declaration for the Responsible Development of Artificial Intelligence* is available online at: <https://www.montrealdeclaration-responsibleai.com>.

²⁰ See the IOSCO-IOSCO report, *The Use of Artificial Intelligence and Machine Learning by Market Intermediaries and Asset Managers* (p. 13), available at: <https://www.iosco.org/library/pubdocs/pdf/IOSCOPD658.pdf>

The following table presents the different principles used by the Canadian financial institutions (FI) and organizations mentioned above.

Canadian FI	OECD	IEEE	FEAT	IOSCO's Fintech Network
Social impact	Inclusive growth, sustainable development and well-being	Well-being		Beneficence Non-maleficence
Equity Non-discrimination Data quality	People-centred values and equity	Human Rights	Fairness (Justifiability & Accuracy and Bias)	
Transparency Explainability/interpretability	Transparency and explainability	Transparency	Transparency	Justice (Accountability & transparency)
Caution Data protection	Robustness, safety and security	Effectiveness		
Human responsibility	Responsibility	Accountability	Accountability (internal & external)	Explainability
		Data Agency		
		Awareness of Misuse	Ethics	Autonomy
Excellent scientist		Competence		

The table shows correspondences and convergences of principles, but also for each list the missing principles. We note that the IEEE (Institute of Electrical and Electronics Engineers) guidance document is the most complete, but also that it includes governance and legal principles that must be linked to higher-level ethical principles. Respect for human rights should not be included, as it is a legal obligation under international law. A competence (IEEE) such as scientific excellence does not belong to a list of ethical principles either and should rather be attached to the principle of accountability, or more satisfactorily included in a good governance scheme as we will do in section 3.3.

By cross-referencing dozens of papers in AI ethics (84 to be precise), some researchers²¹ have shown that there is a convergence on five principles for which there is minimal consensus:

- (1) **transparency and explainability**
- (2) **justice and equity**
- (3) **non-maleficence, safety**
- (4) **responsibility, accountability**
- (5) **privacy**

This cross-referencing approach is interesting to guide the development of responsible AI and it is always better to reduce the number of principles than to multiply them unnecessarily. Moreover, the IOSCO-OICV Fintech Network proposes to include transparency and accountability in the justice principle, would then have only three principles.

The rule for reducing the number of principles is always the same: relevance. However, if certain lists could be reduced, the list of five principles proposed above obliterates ethical issues that are crucial for the deployment of AI in financial markets, such as the principles of autonomy or solidarity, which were put forward by the participants in the consumer consultation. Thus, it is desirable to go beyond the minimal consensus that is established by a cross-reference of institutional documents to take into account the expectations of the public, citizens and consumers.

²¹ Jobin, Anna, Marcello Lenca et Effy Vayena. 2019. "The global landscape of AI ethics guidelines" *Nature Machine Intelligence*, 2019, n°9: 389-399.

4.2. An encompassing framework for AI ethics in finance

The *Montreal Declaration for the Responsible Development of AI* (hereafter “the *Montreal Declaration*”) provides a basis for thinking about AI in finance²². It is an inclusive ethical framework with ten principles and sixty sub-principles for implementation, and is the result of a public, multi-stakeholder and citizen consultation²³, which allowed for the consideration of principles overlooked in the above lists:

MDRAI Principles	Principles covered by the MDRAI
Well-being	<ul style="list-style-type: none"> ○ Kindness ○ Non-malice
Autonomy	<ul style="list-style-type: none"> ○ Consent ○ Explainability ○ Data control
Privacy, protection of intimacy	<ul style="list-style-type: none"> ○ Data protection ○ Data control
Solidarity	<ul style="list-style-type: none"> ○ Inclusive growth
Democracy (control, transparency & interpretability)	<ul style="list-style-type: none"> ○ Transparency ○ Interpretability and explainability
Equity	<ul style="list-style-type: none"> ○ Non-discrimination ○ Data quality ○ Justifiability
Diversity	<ul style="list-style-type: none"> ○ Social inclusion
Prudence	<ul style="list-style-type: none"> ○ Security ○ Robustness ○ Efficiency
Responsibility	<ul style="list-style-type: none"> ○ Accountability ○ Competence
Sustainable development	<ul style="list-style-type: none"> ○ Environmental footprint reduction

²² For a specific analysis of the different ethical frameworks that situate the Montreal Declaration, see Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar. “Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI,” *Berkman Klein Center for Internet & Society*, 2020. It is unfortunate, however, that this study compares the different declarations based on a selection of principles that does not reflect the richness of the principles contained in the Montreal Declaration.

²³ See Abrassart Ch., M-A. Dilhac and N. Voarino (eds.), *Report of the Montreal Declaration*, 2018: https://www.montrealdeclaration-responsibleai.com/files/ugd/ebc3a3_5c89e007e0de440097cef36dcd69c7b0.pdf

Ethical principles are guides for assessing AIS, their impacts, the risks they entail, and the conditions for their social acceptability. Their usefulness depends on their interpretation for application to particular sectors or types of use. It is a matter of moving from the “what” (principles) to the “how” (good practice recommendations, standards and governance mechanisms)²⁴. The following table presents recommendations that have been developed taking into account the principles of the Declaration, on the one hand, and the specificities of several financial sectors on the other²⁵. The list of recommendations is not exhaustive, as the variety of financial sectors, professions, services and contexts in which AI is applied is very large.

Principles	⇒	Applied to financial markets
Well-being <ul style="list-style-type: none"> ○ Kindness ○ Non-malice 		<ul style="list-style-type: none"> ▪ Do not use AI/ML to harm clients (consumers, investors). ▪ Use AI/ML to promote their best interest.
Autonomy <ul style="list-style-type: none"> ○ Consent ○ Data control 		<ul style="list-style-type: none"> ▪ Provide investors and other financial market participants with the best information on how to use AI/ML to help them make good decisions for their clients. ▪ Provide consumers with all relevant information about how the AIS works, along with the evaluation or decision process so that they can give informed consent. ▪ Do not deploy AIS that constrain consumer choices, lifestyles, and beliefs. ▪ Provide exit options for consumers, especially for AIS that influence or monitor behaviour.
Privacy, protection of intimacy <ul style="list-style-type: none"> ○ Data protection ○ Data control 		<ul style="list-style-type: none"> ▪ Protect consumers’ personal data and ensure strict confidentiality. ▪ Ensure that personal data is used for the purposes agreed upon and prevent misuse or duplication. ▪ Allow consumers control over the use of their data (right to data portability).
Solidarity <ul style="list-style-type: none"> ○ Inclusive growth ○ Risk pooling 		<ul style="list-style-type: none"> ▪ AI/ML should help improve risk management and foster market conditions with a more equitable and mutual allocation of individual and collective risks among investors.

²⁴ Cf. e.g. Morley, J., Floridi, L., Kinsey, L. *et al.* “From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices”, *Sci Eng Ethics* 26, 2141-2168 (2020). <https://doi.org/10.1007/s11948-019-00165-5>

²⁵ Cf. Marc-Antoine Dilhac, with Manuel Morales and Rhéia Khalaf, co-authors of this report, developed a similar interpretation of the Declaration in CPA Canada’s contribution to the IOSCO-IOSCO consultation on the use of AI in the narrow field of market intermediaries: “Comments on the use of artificial intelligence and machine learning by market intermediaries and asset managers” (CPA Canada, 2020). The list of recommendations developed in this report is more comprehensive and covers different sectors and clienteles of the Authority, including the insurance sector.

<p>Democracy (control, transparency & interpretability)</p> <ul style="list-style-type: none"> ○ Transparency ○ Interpretability and explainability 	<ul style="list-style-type: none"> ▪ Make codes and training data sets accessible to regulators, and verifiable by competent authorities; ▪ Make codes explainable to regulators and, where appropriate, to customers, according to different standards and levels of explainability.
<p>Equity</p> <ul style="list-style-type: none"> ○ Non-discrimination ○ Justifiability ○ Unbiased competition 	<ul style="list-style-type: none"> ▪ Ensure that MLSs treat similar cases similarly and do not favour one investor over another. ▪ Ensure that the use of AI/ML does not give an unfair advantage to some financial market participants over others. ▪ Prevent biases in data and algorithms that discriminate between investors or in the selection of investments.
<p>Diversity</p> <ul style="list-style-type: none"> ○ Inclusion ○ Social responsibility 	<ul style="list-style-type: none"> ▪ Ensure that AI development environments are inclusive and reflect the diversity of individuals and groups in society. ▪ Ensure diversification in the selection of investments to avoid potentially underperforming investments.
<p>Prudence</p> <ul style="list-style-type: none"> ○ Security ○ Stability ○ Efficiency 	<ul style="list-style-type: none"> ▪ Test before large-scale deployment (simulation, small-scale deployment). ▪ Monitor and implement ongoing upstream (ex-ante) and downstream (ex-post) AIS impact assessment. ▪ Disclose errors and flaws discovered in AIS when they pose a significant threat to investor integrity. ▪ Promote AI-powered financial stewardship.
<p>Responsibility</p> <ul style="list-style-type: none"> ○ Accountability ○ Competence 	<ul style="list-style-type: none"> ▪ Always keep a human in the decision loop when decision-making affects the interests and integrity of consumers. ▪ Ensure that those responsible for the deployment and use of AIS within the financial institution have the appropriate level of skill and knowledge to understand the implications of its use (whether or not its design has been outsourced).
<p>Sustainable development</p>	<ul style="list-style-type: none"> ▪ AI/ML infrastructure should not generate undue environmental impact. ▪ The AIS should enable the regulator to ensure that the environmental commitments of financial institutions are met.

As the table below shows, once an adequate list of principles is in place, a set of recommendations applicable to the field of finance can be developed. Without commenting on each of the principles, we will discuss some aspects of the equity principle, the requirement of transparency and explainability, the respect for autonomy principle and the solidarity principle. Note also that the Declaration covers ESG criteria – they are shown in green in the table with the criterion implemented in parentheses.

4.3. Equity and discrimination

The equity principle is one of the principles of the responsible use of AI in finance, but it is not the only one. However, there is sometimes a tendency for financial institutions to privilege this principle, along with the explainability principle, which we will circle back to, at the expense of other principles²⁶. There is a reason for this: under the pressure of criticism from social science researchers²⁷ and image-destroying newspaper articles, the major concern of companies has been to rectify biases that lead to discrimination against vulnerable groups. This essential criticism has been so prominent that it has considerably narrowed the field of responsible AI thinking, not only in finance. Yet in most cases, the equity principle isn't clearly defined, or when it is defined, it is defined in a circular way, and becomes quite undemanding. How should it be understood?

In the language of market regulators, the equity principle mainly refers to the idea that competition between market participants should not be distorted by illicit schemes such as privileged access to information (insider trading), collusion, fraud or corruption. The market equity principle is the application of the equal opportunity principle to the market and is derived from a more general conception that will allow us to address the issue of algorithmic discrimination. Formally, the equity principle posits that similar cases should be treated similarly – this definition is as old as ancient Greek philosophy. Unequal treatment of similar cases constitutes discrimination. The problem is what is meant by similar cases. Should we consider the case of car insurance for men to be similar to the case of car insurance for women? If so, we shouldn't treat men and women differently in the auto insurance industry²⁸.

²⁶ See Keller, Benno. 2020. *Promoting Responsible Artificial Intelligence in Insurance*. Genève: The Geneva Association.

²⁷ Of particular note is the study by Safiya Umoja Noble (2018), *Algorithms of Oppression: How Search Engines Reinforce Racism*.

²⁸ In Quebec, for example, female drivers pay less than male drivers. See the statistics of the Groupement des assureurs automobiles: <https://gaa.qc.ca/en/statistics/rating-criteria/>

An interesting approach is to combine *equity* and *justification* (and not *algorithmic explanation*, which says nothing about equity). We recommend this approach, which is that of the Monetary Authority of Singapore: “*It is important that AIDA [artificial intelligence and data analytics]-driven decisions do not disadvantage any particular individual or groups of individuals without justification.*”²⁹ This idea is also very old and it is essential to understand that equity is a formal procedure (equal treatment) based on substantial criteria that must be justified. It is therefore necessary to identify *relevant* criteria to establish legitimate differences in treatment³⁰.

This brings us back to the relevance principle presented above and extends its scope. The objectives of AIS must be relevant and the types of data used must also be relevant. Is direct or indirect ethnic data relevant in finance? The answer is clearly no, while it may be yes in the health field. Let’s go further: is ethnic data relevant in health (certain ethnic types are more vulnerable to certain pathologies) relevant for the insurer in the health insurance field? Even if the answer is less clear because a correlation, if not causation, is assumed, it is also negative insofar as this type of data discriminates people not on the basis of health but on the basis of their ethnicity³¹.

And what about data extracted from social networks? Participants in the Authority’s consultation workshops questioned whether it is relevant. To the extent that non-traditional, unregulated data is now being mined by AIS, there is an urgent need for financial institutions, lawmakers, and the regulator to determine the types of data that are reasonable, justifiable, and fair to use with AI in finance.

This is especially important because financial institutions may not always be aware of the transformation of the data sets they use; they sometimes consider the data used with AI to be the same as the data used without AI. In addition to the fact that traditional data collection with AI can be intrusive and privacy-invasive, it is important to be mindful of the fact that new types of data are being mined: in the case of driving assessment, the only data that could be used to assess driving was ex-post reports of driving incidents, accidents, or lack thereof. With AIS, data is collected in real time (in the present) thanks to sensors embedded in vehicles or other types of connected objects. This data is analyzed to provide predictions, and therefore to sanction behaviours and not the result of that behaviour (the damage created), which can be rightly perceived as discrimination: similar cases (driving without an accident) are not treated in a similar way (modulation of the insurance policy).

²⁹ *Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of AI and Data Analytics* (Monetary Authority of Singapore 2018).

³⁰ See Benno Keller, op. cit. p.11.

³¹ On actuarial fairness, see Landes, Xavier. 2015. « How Fair Is Actuarial Fairness? ». *Journal of Business Ethics*, Vol. 128, No. 3: 519-533.

4.4. Transparency, explainability and justifiability

One of the principles that one encounters in all the publications on the ethical framework and regulation of AI, as well as in the various public consultation reports, is that of the transparency and explainability of decisions³². These principles are always associated, but it should be noted that they are two principles that should be treated separately, unless one makes transparency a synonym of explainability: AIS, or explicable algorithmic decisions, are transparent. But this would be redundant and the use of the term transparent would not be rigorous. If explainability requires a certain level of disclosure of the parameters of the decision, the threshold of disclosure can be very low. As for the principle of transparency, it does not require explainability but only the disclosure of information, such as access to the code of the algorithms, their model and the data used. The two principles therefore do not completely overlap, contrary to what the literature on the subject suggests.

The scope of the transparency principle depends on the beneficiary:

- **Consumers** should have access to information about the decision-making process affecting them. This includes whether they interact with an AIS (e.g., a conversational agent), whether a decision about them was made by an AIS or determined based on a recommendation from an AIS, how the AIS is used in the decision-making, who (which service) made the human decision based on the advanced analytics, what parameters and data types the AIS uses, and what personal data was used. This does not include the right to access the algorithm code, or third-party data used by AIS.
 - The **financial institution** deploying an AIS that it did not develop in-house needs to know exactly how the AIS works, what its objectives are, and what results are expected and anticipated. This may include the right to have access to the code in some cases but not necessarily; it does, however, include the right to have access to the mathematical model and databases used to train the algorithms. Without this information, the financial institution is not able to make an informed decision regarding the deployment of AIS, nor is it able to respond to customer requests for explanation and justification.
- The **regulatory authority**, its own departments or authorized supervisory bodies³³ should have access to all of the information mentioned above, as well as the AIS code.

³² Reed Chris. 2018. « How should we regulate artificial intelligence? ». *Philosophical Transactions A of the Royal Society*: 1-12. <https://doi.org/10.1098/rsta.2017.0360>

³³ For various reasons, the regulator may not want to exercise this control. But such control must be exercised, if not by the regulator, by an authorized body. Consumers express this kind of expectation.

Transparency is a condition for interpreting AIS results, but it is not a substitute for explanation or justification.

Even when AIS are transparent, they can remain opaque. This is a paradox in appearance only. Access to the mathematical model and code of an AIS using deep learning techniques does not always allow one to understand the results generated, because the principle of deep learning is precisely to let the system determine its method for achieving the type of result expected. This is why these AIS are called “black boxes”³⁴. However, from an ethical and legal point of view, any decision affecting a person must be explained and it is not enough to delegate the decision to a machine to exempt oneself from the duty to explain³⁵. The judgment of the Artificial Intelligence Committee of the British House of Lords is clear:

« it is not acceptable to deploy any artificial intelligence system which could have a substantial impact on an individual’s life, unless it can generate a full and satisfactory explanation for the decisions it will take. »³⁶

But if we cannot explain the results of AIS that function as black boxes, should we maintain this kind of obligation? To answer this question, it should first be noted that the inability to explain the operation or results of an AIS would invalidate its deployment. Yet invalidation is not the only option, and the complex problem of explainability forces regulators to develop a complex and nuanced

response. The more precise the predictions of AISs, the less explicable or interpretable their operation is even for their developers. Therefore, the pros and cons of explainability must be weighed and the principle of explainability must be

balanced against the principle of robustness or safety³⁷.

Although the efforts of AIS developers to make models explainable help limit the black box phenomenon and several techniques exist to make AIS explainable, one can make the assumption that many AIS with significant benefit to financial institutions and consumers will remain black boxes of sorts. Is this something that should be resolved?



³⁴ See Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge: Harvard University Press.

³⁵ See for example Goodman B, Seth Flaxman S. 2017. « European Union Regulations on Algorithmic Decision-making and a ‘Right to Explanation’ ». *AI Magazine*, Vol 38 (3): 1-9. <https://arxiv.org/abs/1606.08813>

³⁶ UK House of Lords Artificial Intelligence Committee. 2018. *AI in the UK: ready, willing and able?* [7]. Londres: para 105 [cited by Reed Chris (2018)].

³⁷ Lundberg, Scott M., et Su-In Lee. 2017. “A Unified Approach to Interpreting Model Predictions”, *31st Conference on Neural Information Processing Systems*: 1-10. <https://arxiv.org/pdf/1705.07874.pdf>

The answer lies in the distinction between explainability and justifiability. Even if the result of an AIS is not explainable, financial market participants must be able to justify the use of that result. In other words, the responsibility for the use of AIS results lies with the individuals or legal entities and this responsibility consists in assuming the use of algorithmic results. Justifying an algorithmic result means showing that if a human, with the same data, had had to make a decision or a recommendation, they would have been similar or identical to those of the AIS. This is precisely what the Declaration states, “Justification consists of outlining the most important factors and parameters of the decision and should be similar to the justifications that would be required of a human being making the same type of decision.”³⁸ In addition, for AIS that make decisions, financial institutions should explain and justify their models to the regulator according to a practice established by the Basel Accords for bank risk management.



4.5. Autonomy and nudging AIS

It should be noted that some principles advanced by the *Montreal Declaration* are neglected or even completely ignored in the other declarations. This is the case for the respect for autonomy principle and the solidarity principle of. Let's start with autonomy. As the various consultations conducted for the *Montreal Declaration* or for the Authority have shown, this principle meets pressing public expectations. This principle is also one of the pillars of applied ethics, particularly medical ethics³⁹, and is based on the philosophy of the Enlightenment that underlies our democratic practices. Neglecting this principle can result in serious harm to one of the fundamental interests of people, and therefore of consumers.

Taking into account the respect of autonomy, we can evaluate in a nuanced way the AIS in finance that exert small pressures (nudging) on consumers⁴⁰. This is the case of AIS in the insurance field, which evaluate the behaviour of policyholders and encourage them to adopt certain behaviours considered to be healthier or more prudent. A distinction must be made between insurance 53dvent53es that are problematic and those that are less so, such as automobile insurance, but in general, the use of these AIS is legitimate under certain conditions:

³⁸ See Principle 5. Democratic Participation, *Montreal Declaration* (2018).

³⁹ This is one of the four or five authoritative principles in the field of medical ethics and bioethics. This principle is the basis of the notion of free and informed consent as formulated in the World Medical Association's Declaration of Helsinki ("Ethical Principles for Medical Research Involving Human Subjects", 1964).

⁴⁰ See the classic work of Thaler, R. H. and Cass R. Sunstein (2018), *Nudge: Improving Decisions about Health, Wealth, and Happiness*.

- a. AIS should promote consumer well-being by helping them manage certain risks for which they are otherwise covered;
- b. Consumers must agree with the financial institution's objective of the nudging AIS; in other words, they must share the objective of reducing behavioural risk;
- c. Consumers must consent to be observed in a limited way, which infringes on their privacy;
- d. Because of (b) and (c), the financial institution must offer options without nudging AIS;
- e. Because of (a) (b) (c) and (d), consumer refusal to adopt the nudging AIS should not cause them to incur a sanction by the financial institution.
- f. Because of (a), bad behaviour that is not prohibited by law and has no proven effect (damage, illness) should not be sanctioned.

Remember that reward mechanisms are simply the flip side of penalty mechanisms: if some consumers receive discounts because of their behaviour or willingness to use the AIS, it likely means that others pay more for their service.

One might ask whether it is still worthwhile for a financial institution to implement nudging AIS if the consumer who adopts it isn't rewarded and the consumer who doesn't adopt it and engages in deemed risky behaviour isn't sanctioned. Yet the answer is obvious in the context of an ethical deployment of AI in finance: the nudging AIS must be designed to promote the good (health, safety) of consumers (a) and, out of respect for their autonomy, the institution should instead simply offer the AIS as a resource to promote their own good. If consumers adopt it, the financial institution wins because consumers who adopt this type of system will tend to improve their behaviour since it is for their own good. Consumers will know that the AIS is all about

promoting their good and that the consequences of their more virtuous behaviour are also good for their financial institution. Each party wins for different but congruent reasons⁴¹.

These principles don't apply with the same force in different insurance sectors. For example, virtuous incentive mechanisms are less problematic in the area of automobile insurance than in the area of health insurance. In automobile insurance, the incentive to drive well achieves a goal that is legal, not merely ethical. Driving safely, respecting traffic laws, including speed limits, is a legal obligation and all policyholders have the same obligation. The AIS of virtuous driving incentives and verification encourage drivers to comply with this legal

⁴¹ Marc-Antoine Dilhac, co-author of this report, developed this approach in the white paper on the use of AI in the insurance sector: *Artificial Intelligence, Solidarity and Insurance in Europe and Canada* (Institute of Technology for People, 2020) - the white paper was developed in partnership with OBVIA as well as key insurance companies (CNP, SwissLife, Allianz, Intact).

obligation. The case is different in the field of health, where lifestyle behaviours are not subject to legal obligations – except in very exceptional situations defined by health authorities. In this

area, AIS can seriously impinge on the autonomy of individuals, and their deployment should be subject to strict supervision.

4.6. Solidarity

The solidarity principle doesn't appear in any of the major declarations mentioned. Yet it is crucial for the stability of financial markets. This principle points in two directions specified in the Declaration.

First, solidarity means that we must always give priority to **collaboration between humans and machines**, rather than the replacement of the former by the latter. This reinforces the responsibility principle to keep “a human in the loop” of decision-making. This general solidarity principle can be applied to all sectors of finance and has concrete effects on the governance of AI in financial institutions.

The other meaning of solidarity, which is general but applies particularly to the insurance sector, is that of **risk pooling**⁴². Solidarity is the moral expression of the need to pool risks from which the insurance system emerged. This principle requires that risks be artificially distributed equally within a given population. A loss will happen to some members of that population, but the probability of it happening is considered equal for all members, so that all have an interest in pooling resources to be compensated in case of a loss. But suppose we could know which members would suffer the loss; would the other members have the same interest in pooling their resources? Intuitively, unless we admit that individuals are perfectly altruistic, the answer is no.

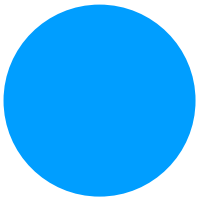


⁴² See *Artificial Intelligence, Solidarity and Insurance in Europe and Canada* (Institute of Technology for People, 2020).

Insurers have an interest in predicting risks and adjusting their insurance policy according to risk profiles, which is known as segmentation; for their part, policyholders can sometimes be happy to pay less than other policyholders – this is the main selling point, and the main reason for acceptance. However, the emergence of AI techniques to improve risk prediction and knowledge of risk behaviours raises a major ethical issue: what is the future of solidarity in the insurance sector⁴³? And what is the future of insurance itself?⁴⁴ Nevertheless, not all damages can be attributed to the risk behaviours of policyholders, nor can they be predicted with certainty, and segmentation, although more precise with the deployment of AIS, cannot be pushed to the point of hyper-personalization of risk profiles.

⁴³ Charpentier Arthur, Michel M. Denuit et Romuald Elie. 2015. « Segmentation et Mutualisation, les deux faces d'une même pièce ». *Risques. Les Cahiers de l'assurance*, n°103 : 19-23.

⁴⁴ Daniel, Jean-Pierre. 2016. « Et si l'assurance disparaissait ? », *Risques. Les Cahiers de l'assurance*, n° 108: 123-127.



5. AI governance mechanisms

Having determined the objectives sought through the adoption and use of an AIS and having established the principles of ethical compliance, one must examine the governance mechanisms that enable ethical compliance to be achieved. Without governance mechanisms, either recommended by the regulator or imposed by law, ethical compliance remains a theoretical idea and the risks to consumers, financial market participants and the stability of those markets cannot be eliminated or even reduced.

Most of these mechanisms can be built on top of existing structures in financial institutions. Effective reorganization is less costly and more beneficial than creating new bodies, although sometimes new bodies are still needed.

5.1. Ethical conduct

Financial institutions should establish a code of ethics for professionals developing in-house AIS and using AIS in their work⁴⁵. This code of ethics should specify the ethical principles of AI development and use, provide an accessible interpretation of these principles, contain a repertoire of good practices, and reiterate professional virtues and obligations in general.

This proposal was evident in the public and multi-stakeholder deliberations in which participants also mentioned a code of ethics. While it may be felt that professional orders and associations, in the case of engineers for example, should update codes of ethics to take into account the new ethical obligations related to the development and use of AI, the first step is to adapt codes of ethics for which financial institutions have more leeway.

⁴⁵ Note that this mechanism corresponds to the *Ethics* principle of the Monetary Authority of Singapore's proposal. In reality, it is not strictly speaking a principle but a governance mechanism.



5.2. Impact assessment and audit

The deployment of AIS in finance requires an impact assessment on customers and employees. These assessments should be implemented internally (within the financial institution) and externally by the competent authority.

Impact assessment⁴⁶ involves ensuring *ex ante* that the AIS is designed in accordance with the principles of AI ethics (ethical compliance) and that the foreseeable impact of its deployment in the organization does not violate ethical principles (and human rights).

These *ex-ante* evaluations must be followed by *ex-post* evaluations by comparing the desired effects with the actual consequences. In order to do this, feedback loops must allow the identification of operational anomalies, unforeseen and undesirable effects: internal reports and feedback from clients (users or beneficiaries). It is crucial to

be able to precisely document the consequences of the use of AIS in order to correct a worrisome operation as soon as possible.

Financial institutions should adapt their ethics committees where possible and appropriate or create teams to carry out these controls and impact assessments. In both cases, the teams should be multidisciplinary and this must include the presence of trained ethicists whose specific expertise must be distinguished from legal expertise. The expertise of the ethicists makes it

⁴⁶ See the contribution by Dillon Reisman, Jason Schultz, Kate Crawford et Meredith Whittaker. 2018. "Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability". *AI Now*. New York. <https://ainowinstitute.org/aiareport2018.pdf>

possible to identify issues that are not covered by the law and to anticipate new risks.

While internal control is essential, given the risks of misuse or poorly controlled deployment of AI, external control must also be possible. Rather than creating a central agency dedicated to impact control, it seems more practical and realistic in the short and medium term to rely on existing structures such as financial market regulators.

This external oversight implies that the regulator can request audits of the AIS used and an

assessment of its deployment in the financial institution. The regulator must be able to ensure, for example, that the financial institution has appropriate governance mechanisms in place and that the financial institution is able to provide the supervisory body (not necessarily the regulator⁴⁷) with access to the code and databases and that it is able to identify risks and correct malfunctions. This is the meaning of the principle of transparency, which does not mean that this information is available to the general public, but first to the competent authorities.

5.3. Certification

To control the quality of AIS upstream, certification is a popular mechanism for consumers and researchers. Certification, which is the counterpart of auditing, should not be limited to the safety and robustness of AIS, but should validate the ethical compliance of the product. The ethical compliance principles developed above from the Declaration, and other reference documents for AI governance, provide a solid foundation for certification.

AIS certification is an international issue both to recognize credible institutions in the certification offer and to establish the content of the certification: technical and ethical standards. Who will have the authority to certify an AIS or a financial product using AI? Some researchers proposed the creation of a Certification Agency⁴⁸. Although attractive, this idea seems difficult to implement. It is again appropriate to turn to existing organizations that have developed ethical and legal expertise in AI and have national and international credibility. Regulatory authorities can play a role in catalyzing the supply of certification and providing guidance.



⁴⁷ Given the magnitude of the task, it is not clear that regulators will want to move in this direction. However, it is conceivable that if the regulator receives a number of consumer complaints about an AIS product, it will want to ask the financial institution to explain its system and justify its use. The burden of proof would then be on the financial institution.

⁴⁸ Matthew U. Scherer. 2016. « Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, And Strategies », *Harvard Journal of Law & Technology* 29 (2): 353-400.

5.4. Accountability and Recourse

The opacity of AI use and the serious consequences of making a bad algorithmic decision (or one based on an algorithmic recommendation) are of great concern to consumers. The opacity that worries is less about the algorithm codes (if they were accessible and transparent, it wouldn't change the perception of opacity) and more about the decision-making processes. But it is difficult for customers to identify an algorithmic decision, to ask for explanations and, if necessary, to make a complaint. Since it is recognized that financial institutions that use AI must be accountable to their customers for the decision-making processes using AI and offer a justification for decisions affecting customers, it is appropriate that they have a body dedicated to the accountability of the uses of AI across the different departments or sectors of their organization. While existing structures should be used whenever possible to meet new challenges, it seems that in this case, given the complexity of the decision-making processes, a new centralized structure that has a global view of the organization and services is relevant to meet the demands for justification and to distribute responsibilities internally.



Access to a reasoned **justification** for algorithmic decisions is a condition for exercising the **right to redress** to review the decision or obtain redress. Information for exercising consumer redress must

be readily available, but the processes for handling disputes for review or redress must also be readily available.

In the event of a dispute, it is desirable to use flexible and rapid dispute resolution mechanisms such as mediation and, where appropriate, arbitration. The regulator can encourage the use of such mechanisms and ensure that they are easily accessible to consumers. Finally, it should be added that financial institutions and other financial market players most often use externally developed AIS (algorithmic solutions and databases). This supply of AI solutions can also give rise to complaints from these users (financial

institutions and other financial market players), when the system does not achieve its objectives, is not secure, makes decisions or recommendations that are difficult to justify and are massively contested by their clients, etc. Disputes between companies developing AI and those using their solutions are multiplying⁴⁹ and mediation and arbitration mechanisms seem to be the most efficient way to settle these disputes quickly and thus cause the least damage to the clients of financial institutions and market players.

5.5. Training and awareness

One proposal often comes up in the AI consultations: training. There is a consensus on the need to raise awareness of the ethical challenges of AI. This proposal is aimed at AIS developers in finance, financial market professionals who directly or indirectly use AIS, along with consumers. But the objectives are not the same for each targeted audience.

- a. **Developers:** The training reminds us that what can be done with AI is not in itself desirable, ethical and acceptable. What can be done doesn't necessarily have to be done. This training is beyond the control of financial institutions and the regulator, but they can encourage training institutions (e.g., universities) and professional bodies to offer it.
- b. **Financial market professionals:** This training complements adherence to a code of ethics and conduct in financial institutions. It is important to ensure that professionals understand the duties and obligations specified in the code of ethics. This training can be provided internally by the ethicists and lawyers who serve on the ethics committee, as recommended in section 3.3.1. It can also be outsourced. In this case, it must also be ensured that the training is provided by competent ethicists and lawyers. It is up to the financial institutions to set up such training, and it is up to the regulator to ensure that it is planned and delivered by the financial institution.
- c. **Consumers:** The goal is to both demystify AI and raise awareness of the risks of AIS use by financial institutions. This empowers consumers to make an informed choice, be able to identify a problem, ask for justification, or seek redress. The Authority can play an important role in educating the public about the ethical and legal implications of AI use.

⁴⁹ While not specifically about the use of AI but rather digital in general, Ethan Katsh and Orna Rabinovich-Einy (2017) provides a relevant analysis for understanding the evolution of litigation involving new digital technologies, and legal solutions.

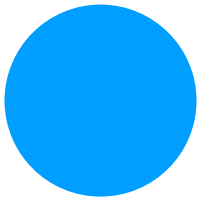
5.6. Consumer consultation

Consulting with consumers is a mechanism that is increasingly used to ensure **the relevance and acceptability of AIS**, but also their ethical compliance upstream. This can be seen in the medical field in particular, where multidisciplinary and multi-stakeholder panels (physicians, patient partners, patients' families) are set up to test AIS and refine their objectives and deployment methods⁵⁰. As in medicine, in finance, consumer consultation should not be conceived as a marketing strategy to identify the product features that will appeal most to consumers. It must be conceived as the integration of a different expertise that allows the identification of ethical risks of the development and deployment of an AIS that professionals engaged in the development and marketing of an AI product are likely to overlook or even ignore.



The integration of consumer expertise should take the form of deliberative consultation, i.e., discussion workshops where participants exchange arguments and strive to reach a consensus – this is the model for the consultation conducted for the Authority in this report. Opinion surveys have their merits, but they capture only snapshot opinions, raw preferences, and sometimes just emotions. Deliberative consultations can be conducted internally when the financial institution is large enough to provide this type of oversight mechanism, or they can be outsourced, which has the advantage of independence of consultation.

⁵⁰ A remarkable example is the development of Glass, an emotion detection device for autistic people, by a team at Stanford: <http://autismglass.stanford.edu>.



Appendix 1. The use of AI in finance in Quebec and Canada

It is not easy to map the development and use of AI by financial institutions in Quebec (and in Canada). The information is difficult to collect if one doesn't take the trouble to consult the financial actors themselves. The landscape of AI use that we present here was facilitated by a series of interviews with various stakeholders from financial institutions, asset managers and fintechs in Quebec and Canada that we have grouped into 4 main groups⁵¹:

- Portfolio Managers
- Fintech
- Insurers
- Deposit-taking institutions

We will begin with the commonalities between the different sectors and then examine the specifics of each sector. One consideration reported by all the stakeholders we met is the availability of data. The integration of AI techniques requires a marked improvement in information management. It is not enough to have access to data, but it is also necessary to ensure its quality, its “availability, integrity and confidentiality” (AIC) in a format conducive to its use. Building data warehouses has been one of the most important efforts of financial institutions and asset managers in recent years, and is still the main obstacle to the development of fintech firms. The creation of units dedicated to the maintenance and development of databases and their processing

⁵¹ Apart from the purely factual elements, this section only reflects our understanding of the use of AI as it emerges from our exchanges with the financial institutions that agreed to meet with us. We do not report on any specific statements. We thank the financial institutions, organizations and firms that agreed to answer our questions (PSP, Intact Insurance, Desjardins Insurance, BNC, CDPQ, Manulife, Fintech Cadence).

has required considerable financial and human resource investments. Although significant progress has been made by many, the consensus is that additional resources for information processing will be required in the coming years.

These large digital databases have led to increased cybersecurity risks related to data security. The various speakers mentioned three major reasons for the massive investments committed to cybersecurity, namely compliance with privacy regulations and existing contracts, ethics or respect for their customers, and reputational risk from security breaches.

The second point in common for all stakeholders is the digitization of repetitive processes with low added value. The objective is not necessarily to reduce labour costs, but rather to free the workforce from repetitive tasks so that they can spend more time on high-value-added mandates. On the contrary, the organizations we interviewed all mentioned that employment in sectors subject to greater automation has generally remained stable or even increased. In some cases, the digitization of certain processes requires AI. The result has been a significant improvement in the operational efficiency of Canadian and Quebec financial institutions and asset managers.

The third point in common is the improvement of the customer experience. The consumption of financial products is now largely done online and less and less in person. This trend has forced financial institutions to significantly increase their digital solution development capabilities to meet customer expectations while enabling increased operational efficiency.

As consumers use digital financial services more, they expect more, especially compared to the services they get from the major consumer Internet companies. Digital experience leaders are continually raising the bar on personalization, to the point where they sometimes anticipate customer needs before the customer is aware of them, and offer tailored services at the right time, via the right channel. Financial institutions are therefore forced to follow this trend.

Finally, the digitization of processes and the use of AI requires very high financial and human resource investments. The barriers to entry are therefore high and could put small and medium-sized groups at a disadvantage across Canada. For many applications, the development costs are the same regardless of the size of the company. In this context, larger institutions have a greater advantage. The challenge for smaller financial institutions is to be agile and focused in their development. Although some applications are developed with external suppliers, most solutions are developed by in-house teams, hence the importance of human resources.

Asset managers

For many of the major asset managers, the last few years have been devoted to developing information systems prior to using statistical learning techniques. Indeed, the creation of units dedicated to the

maintenance and development of databases and their processing have required considerable financial and human resource investments.

To date, AI investments have primarily affected two main areas:

- Digitization of repetitive processes with low added value
- Augmented intelligence

Augmented intelligence extends and supports human intelligence with artificial intelligence. The technology can analyze large amounts of data in seconds and attempts to find patterns and correlations through machine learning. The results of this data analysis ultimately serve as a basis for humans to speed up and simplify the decision-making process for future actions. Augmented intelligence is not there to replace human intelligence, but to enhance it. In most cases, the goal of these new augmented intelligence applications is not to increase investment performance, but rather to enable portfolio managers to do their jobs better and make their decisions with more confidence. Therefore, most final portfolio management decisions are still made in the traditional manner.

The use of technology solutions is expected to increase significantly over the next few years. On the one hand, the acceptance of AI techniques within organizations argues for increased use in the future. On the other hand, the competition for superior efficiency will push organizations towards increased use of AI techniques. In addition, the widespread use of AI in financial markets may force some asset managers to compete intensely in certain asset classes where they can compete and adopt passive management in other asset classes.

The fintech sector

The term “fintech” is used to refer to a company that works in the field of financial technology. Fintechs are generally information technology start-ups that offer financial solutions, often in areas where large financial institutions are either not very innovative or lagging in the adoption of new technologies.

The challenges for fintech start-ups mainly include:

- Access to venture capital
- The demand for their products
- Access to financial data

The fintech business model is based on the ability to identify and respond to a need not covered by large financial institutions. In Canada, this agility is often confronted with the realities of the domestic financial market where there are large players sharing a relatively small market. These large players are established, and innovation is often done internally. In this context, fintech companies often need to adopt a more global approach that is not solely focused on the domestic market.

Despite this, the fintech ecosystem remains vibrant and attracts a lot of young talent. Fintechs occupy most sub-sectors of the financial industry, but with an emphasis on retail applications, such as payment systems or decision support systems like price comparators.

In addition to the difficulties of the Canadian market, the biggest challenge for fintechs remains access to financial information. Without access to information, it becomes difficult, if not impossible, to develop or propose innovative solutions. The large financial groups that have invested significant amounts in their data warehouses could lose their competitive advantages if access were to be expanded, especially in the face of the big global players who could both use their development strength and apply it to the Canadian context. The growth of the fintech industry in Canada therefore depends in part on increased availability of information sources.

The insurance sector

AI technologies are rapidly expanding in the insurance field. This industry has significant information technology resources especially due to product pricing and claims evaluation.

In recent years, the industry has expanded its digital capabilities with new technological opportunities. Artificial intelligence and its related technologies are having a major impact on all aspects of the insurance industry, from distribution to underwriting to pricing to claims. Advanced technologies and data are already affecting distribution and underwriting, with policies being priced, purchased and consolidated in near real time.

These models use internal data as well as a large set of external data accessible through data providers. The availability of information is of considerable importance to insurers. The more information an insurer has access to, the more it allows them to optimize their pricing policies and claim evaluations.

As a result, the insurance industry is increasingly incorporating cutting-edge AI techniques into:

- Subscription/pricing
- Claims assessment
- Security
- Automation and digitization of processes
- Smart devices

In some segments, price competition is intensifying and thin margins have become the norm, while in other segments, unique insurance offerings are enabling margin expansion and differentiation. Insurance products based on dynamic use and adapted to the behavior of individual consumers, especially as a consequence of the development of connected objects, are becoming more and more prevalent.

AI uses are found in customer service applications including claims, fraud, and automated medical service systems that increasingly manage interactions with policyholders. The turnaround time for the resolution of many claims has been significantly reduced and simplified. Claims management with human intervention is increasingly focused on a few areas: complex and unusual claims, contested claims where human interaction and negotiation is enhanced by data-driven analytics and insights, and claims related to systemic issues.

AI and technology are making insurers more focused on monitoring, preventing and mitigating risks. The Internet of things and new data sources are being used to monitor risks and trigger interventions when factors exceed AI-defined thresholds.

Because the insurance business in Canada is fragmented and the market size is relatively small, insurers face internal competition and are vulnerable to international competition. This is prompting Canadian insurers to take a proactive approach to AI techniques and process digitization to increase productivity, reduce costs and have a renewed offering.

Contrary to what is often feared, the increased automation of tasks through the deployment of AI does not generally lead to job losses but is rather seen as a positive factor. In addition, the increased use of AI and the digitization of processes for the entire economy opens up new markets for the insurance industry such as cybersecurity risks or insurance on demand.

Deposit-taking institutions

The uses of AI in deposit-taking institutions are being facilitated by lower data storage and processing costs, increased access and connectivity for all, and rapid advances in AI-related technologies. These technologies can lead to higher automation and, when deployed after risk control, can improve human decision-making in terms of speed and accuracy. The potential for value creation in the banking industry is one of the largest among all industries.

The three main channels through which banks can use AI are the front office (conversational banking), middle office (fraud detection and risk management) and back office (underwriting). Front and middle office AI applications offer the greatest opportunity for cost reduction in digital banking. In an environment where traditional intermediation margins are gradually eroding, controlling operating costs is of considerable importance. Further digitization of processes and increased use of AI techniques are expected to accelerate in the coming years and contribute significantly to improving the operational efficiency of deposit-taking institutions.

AI technologies can help increase revenue through increased personalization of customer services, reduced error rates, better credit report analysis, and improved resource utilization. They also enable new and previously unrealized opportunities through an improved ability to process and generate insights from vast amounts of data from their customers' experiences and the financial markets.

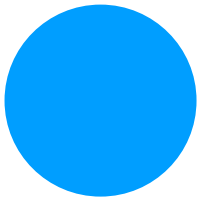
Banks are leveraging algorithms on the front end to facilitate customer identification and authentication, to mimic live employees via conversational “chatbots” and voice assistants, and to improve customer interactions, such as in the case of appointment scheduling.

Many AI applications seek to improve the customer experience. The objective is to make it easy to use new digital services, to improve information and decision-making. It appears that the term augmented intelligence is often preferred to artificial intelligence. Financial institutions are also offering online savings solutions at a lower cost similar to robot-advisors. Once again, the customer experience becomes paramount for these new avenues of financial management to attract and retain users. The benefits are enormous in terms of operational efficiency but also for the personalization of services.

AI is also being implemented by banks within middle-office functions to assess risk, detect and prevent payment fraud, improve anti-money laundering (AML) processes, and perform regulatory know-your-customer (KYC) checks.

According to our meetings, deposit-taking institutions are facing four current trends:

- Growing customer expectations as digital banking adoption increases.
- A growing number of banking executives are taking a holistic approach to deploying AI solutions. The argument of competitiveness takes its stake both in terms of profitability and product differentiation.
- Digital ecosystems are disintermediating traditional financial services. Advances such as blockchain in the payments system are forcing deposit-taking institutions to increase their investments in information technology.
- Technology giants are entering financial services as a logical extension of their core business models.



Appendix 2. Use cases of AI in finance

Although the literature on AI is abundant, it is not always easy to identify documented use cases, especially in finance. By distinguishing four main functions of AI in finance (valuation, incentive, optimization and information), it was possible to highlight ten representative use cases. Five of these use cases inspired the scenarios submitted for collective reflection during the deliberations with citizens and consumers.

Below is the summary table of the ten use cases and then the detailed presentation of use cases 5 to 10 (use cases 1 to 4 were presented earlier, in section 1).



Context	Field	Data and inputs	AI Model	Tasks and results
Mastercard : Decision Intelligence	Security	Transactions, customer profile, CRM, metadata	Knowledge representation, neural network, sequential model, predictive models	Decision and fraud detection service
Lenddo	Credit	Credit data, digital profile from third parties	Predictive models, knowledge representation, computer vision	Data analysis for credit investigation purposes
Desjardins : Ajusto	Insurance	Telematics data, customer profile, metadata, geolocation	Telematics, machine learning (supervised/unsupervised), knowledge representation, predictive models	Driving analysis for personalized performance feedback
Truvalue Labs	Stock Markets	Corporate disclosure, public reporting, historical market data, news, social media	Sentiment analysis, NLP, NLU, OCR, computer vision,	Automated analysis and valuation of data in responsible investment
Kavout	Stock Markets	Market data, public balance sheet, alternative data	Machine learning, predictive analytics, machine learning, deep learning, reinforcement learning	Stock performance analysis, identify stocks likely to outperform
Deep Knowledge Venture : VITAL	Stock Markets	Due diligence data, database	Predictive analysis, DL, NLP, NLU	Risk analysis, investment assistance
Finn AI and Glia	Service	Interaction data, sentiment analysis	NLU, NLP, reinforcement learning	Customer service, optimization of employee time
John Hancock : Vitality Program	Insurance	Biometric data, clinical demographics, customer profiles, metadata, geolocation	Predictive analysis, knowledge representation, machine learning (supervised/unsupervised)	Health monitoring, insurance policy optimization, behavioural reinforcement
IBM : Watson	Security	Transactions, customer profile, CRM, metadata, structured/unstructured data	Cognitive security, machine learning, deep learning, NLP,	Threat analysis and detection, real-time information
Fligoo and Broadbridges	Service	Transactions, customer profile, CRM, metadata, demographic data	Predictive analysis, closed loop feedback, knowledge representation	CRM Optimization

Case 5. Mastercard *Decision Intelligence*: fraud detection

Field: Cybersecurity

Primary use: Optimization

Secondary use: Consulting and information

Context

In 2016, *Mastercard* announced the launch of *Decision Intelligence*, a decision assessment and fraud detection service. The service uses AI to help financial institutions improve the accuracy of genuine transaction approvals and reduce the number of “wrongful denials⁵²” in real time.

Description and objectives of the system

Decision Intelligence takes a broader view than traditional scoring systems in the way it evaluates, scores and learns from each transaction. This scoring then allows the card issuer to apply the new rules it infers from its learning to the next transaction. *Decision Intelligence* looks at how an account is used by its holder to discern normal and abnormal purchasing behaviour by the holder. The technology would be able to make these detections based on a variety of data sources, including: risk assessment, geolocation, information about the merchants involved in transactions, payment device data, time of day and nature of purchases.

System benefits

For banks, the use of this technology reduces their operating costs and increases their revenues. For retailers, it represents a risk reduction. It is estimated that for every \$1 lost to fraud, the recovery cost for financial institutions is nearly \$2.92. In addition to being easy to use, the system improves the customer experience for cardholders and increases their loyalty. The information the system uses could also be used to react much more quickly to problems, potentially reducing operational expenses such as chargebacks.

Technology used

Decision Intelligence is based on real-time transactional data and external data such as anonymous and aggregated customer information, in addition to geolocation information. The latter helps contextualize the nature of a purchase in a given area and determine whether it is “normal” or “credible”. It also helps detect patterns of fraudulent activity.

Decision Intelligence also uses natural language processing (NLP)⁵³ to interpret handwriting and textual data. In this particular case, NLP algorithms can determine links between names, first names and groups of people. Thus, activities or individuals using borrowed, modified names and pseudonyms can be identified.

⁵² Wrongful Refusal: When a transaction is refused because it is falsely deemed fraudulent by the system.

⁵³ Natural Language Processing (NLP).

Case 6. *Ajusto* from Desjardins: telematics for automobile insurance

Field: Insurance (damage)

Primary use: Incentive

Secondary Use: Evaluation

Context

Ajusto is an optional program of the Desjardins automobile insurance policy. It is a telematics-based driving assessment application: *Ajusto* analyzes the driving behaviour of subscribing clients, and rewards good driving habits with discounts on the insurance policy. The program has been around since 2013, but in its first iteration, the program provided their members with a device that plugged into the car to monitor their driving. Drivers could go to the program's website to view the machine's report on a virtual dashboard. In 2015, *Desjardins* migrated the *Ajusto* system to a smartphone app. The telematics work no longer requires the company to send a dedicated device: it's now the phone that collects the driving data. The dashboard can be consulted in real time from the application.

Description and objectives of the system

Ajusto analyzes driving habits and provides the driver with a score and personalized feedback on their performance. The score is based on driving habits such as distance travelled and time of day, and different criteria such as: speed, rapid acceleration, hard braking and cell phone distraction.

The program allows the insurer to customize the risk level of the insurance policy based on the consumer's driving behaviour and to offer a fair price. Following the evaluation period, the client can receive a personalized premium that can be advantageous or disadvantageous: a good score indicating safe driving could result in a reduction in the premium of up to 25%. However, if their score indicates more risky driving, their premium could be increased by up to 20%.

System benefits

This type of system has the potential to improve road safety if adopted by a majority of road users, and 76% of *Ajusto* users say they believe the app contributes to this effect. As a result of striving to drive better to take advantage of the discount offered, nearly 75% of respondents said they had improved their driving. They said they were better at obeying speed limits and avoiding hard acceleration, hard braking and turning too quickly. The opportunity to save money while being more careful is a significant benefit to users. In addition, it allows the insurer to offer a price that is representative of the risk that each of its customers represents, and to collect everyone's fair share.

Technology used

Programs like *Ajusto* use telematics to make their assessments, a model that defines "usage-based insurance"⁵⁴. While some programs use external "black box" devices to plug into the OBD-II of automobiles, most now use the information that the smartphone provides to the insurer's published application. The information from the OBD-II can be used to complement the information from the phone. The data collected by this equipment is then modelled and analyzed by the application, which uses it to assess the risk and eventually offer a new premium to the insured.

⁵⁴ Usage Based Insurance (UBI).

Case 7: *Kavout* and investment assistance

Field: Investment

Primary use: Optimization

Secondary use: Consulting and information

Context

The stock market can be difficult to understand for members of the public who want to become their own financial advisor. For this reason, *Kavout* has created an AI-based investment platform designed for investors of all levels, from the experienced to the novice, to advise them and create portfolios based on massive data analysis.

Description and objectives of the system

The *Kavout* AI platform uses machine learning and predictive analytics combined with a quantitative analysis model - called the *Kai* Index - to analyze stock prices and make a short-term identification of top and bottom-performing stocks. The model takes into consideration news items, and information from blogs and social media.

The *Kai* Index incorporates 200 different correlation measures, then assigns a *Kai* score to each stock. A simple stock score - from 1 to 9 - was developed with machine learning. The higher the *Kai* score, the more likely a stock is to outperform the overall market over the next month. *Kavout* uses this approach to create stock portfolios based on *Kai* scores. The best portfolios are composed of stocks with a high *Kai* score.

System benefits

The tool takes emotion out of the equation and uses data analysis to predict the best outcome. The analysis it does through its *Kai* model allows it to make very accurate recommendations about the value of the stocks they are monitoring. From its sources of news, blogs and social media, *Kavout* can provide all the necessary information surrounding a particular company, in addition to the headlines of companies in the same sector. Thus, *Kavout* gives a panoramic view on the market state of certain industries in its portfolio management.

Technology used

Kavout processes millions of data sets daily, and runs models encompassing many traditional and advanced financial engineering methods, such as: regression, classification, deep and reinforcement learning. It also uses natural language processing to analyze the text of blogs and social media that they scan for information. With all of this, it is able to produce predictive scoring to rank stocks. All this amalgamation makes *Kavout* able to combine massive data from the stock market faster and more efficiently than traders to identify and model investment information.

Case 8. *Glia* with Finn AI: conversational agents

Field: Service

Primary use: Consulting and information

Secondary use: Optimization

Context

Conversational agents (CAs), also known as chatbots, are computer programs that simulate a conversation with their interlocutors by emulating the answers that a human would give. CAs respond either by voice command or by text. CAs can be integrated into messaging applications and are increasingly used in the service industry as a first contact with the customer. That's what's happening over at Finn AI, a CA developer that specializes in banking services. They have partnered with Glia, a company whose expertise is in digital customer service.

Description and objectives of the system

A Finn AI CA has been integrated on a Glia customer service platform. This CA was designed to enable Glia's financial institution clients to improve the customer experience and reduce response time to their members, while increasing the efficiency of its services.

Finn AI's CA is pre-trained to understand and support over 500 banking queries. One of the advantages of this type of technology is that it can provide services 24 hours a day. Glia's services have the advantage of allowing its business customers to use multiple CA engines and counter the vendor lock-in phenomenon⁵⁵.

System benefits

The AI usage model that sees the most performance gain is one where humans and machines work together to divide the task, or AI takes over repetitive tasks and humans can undertake more complex tasks without worrying about routine calls.

In banking customer service, 75% of the time is spent on repetitive routine requests. It is precisely the repetitive nature of these interactions that makes CA attractive. The remaining 25% of the time is spent on complex queries that are customer-specific situations requiring direct assistance from a human customer service agent. By relegating CSRs to high volumes of routine calls without complexity, customers expect less, and workers are left with more time to focus on more complex tasks that value their expertise.

Technology used

The first CAs served mainly as an interactive spokesperson for the FAQ of the company it represented. Answers are programmed in advance, with simple rules linked to keywords. These processes have been improved by Natural Language Understanding (NLU) processes from machine learning

⁵⁵ Vendor lock-in is a phenomenon where granting exclusive ownership to the manufacturer of a software or machine forces an exclusivity and a dependency relationship with the user.

Case 9. IBM's *Watson* and the challenge of cybersecurity

Field: Cybersecurity

Primary use: Optimization

Secondary Use: Evaluation

Context

Cybersecurity issues concern all sectors. The financial sector is particularly concerned by the sensitive data it deals with and the uses it makes of it. Insurers are dealing with issues of protection of personally identifiable data. Banks also deal with the need to secure their customers' assets. The malicious use of technology, progressing at the same pace as defensive capabilities, requires these institutions to implement effective prevention protocols to protect themselves against increasingly sophisticated threats. AI's ability to analyze a large amount of data in a short period of time allows it to make predictions and detect risks and suspicious behaviour in real time. AI technologies are therefore a valuable ally for companies working on cybersecurity solutions

Description and objectives of the system

Watson uses cognitive security to deploy its cybersecurity solutions. Cognitive security is an AI application that is based on human thought processes to recognize security threats. Each cognitive AI interaction proactively teaches the model to detect and analyze threats in order to provide guidance to analysts who can take actions based on the insightful information provided by the AI.

System benefits

Thanks to the interactions that have trained it, *Watson* can make high-speed connections between events that could constitute a cybersecurity threat and make recommendations. Analysts can therefore respond more quickly to potential threats, eliminating time-consuming research and analysis tasks. They are then able to make critical decisions and start a defensive response in a hurry. *Watson's* capabilities have helped analysts with their investigations and have reduced the time required from weeks to hours.

Technology used

The *Watson* system uses cognitive computing processes that are supported by machine learning and deep learning algorithms that allow it to become increasingly powerful and intelligent. It also uses natural language processing to deploy its capabilities. With this capability, *Watson* is able to "read" and learn, pulling information from unstructured data sources such as cybersecurity blogs, websites and resource documents. It has pulled information from more than two million cybersecurity documents.

Watson's AI is trained with billions of pieces of structured security information from threat intelligence feeds, security events and related data. As its model improves, the AI becomes increasingly adept at understanding and identifying potential cybersecurity risks. AI gathers knowledge and uses reasoning to identify relationships between different threats it identifies between malicious records and IP addresses.

Case 10. Fligoo and Broadbridge: Personalized customer service

Field: Service

Primary use: Advice and information

Secondary use: Incentive

Context

Broadbridges Financial Solutions - a company that provides communication services and technology solutions to investment firms, broker-dealers, mutual funds and issuers - has partnered with *Fligoo* to create a service to help banks and wealth management firms use and leverage the data they have on their clients. Using *Fligoo's* AI expertise, together they will develop a proprietary software suite that will help banks and wealth managers make accurate predictions about their clients' needs, tailoring a customized client-investor experience.

Description and objectives of the system

Among the arsenal that the duo offers, the first to be developed will serve, through a data-driven approach, to optimize the service relationship between financial advisors and the investors who consult them. The services rendered and the way they are rendered are tailored to the client's needs. Since investors have an expectation to be well understood by their advisors, the use of AI allows these advisors to anticipate the needs of their clients' investment trajectory. The program will also be able to predict which clients are in need of optimizing loan parameters and terms based on their situation. They can be informed which types of loans and financial products can help them save money, or which one is best suited to their debtor profile. The predictive model can even interpret an investor's specific needs. For example, if the program interprets that they are considering a financing strategy for their child's education, the program will be able to make personalized recommendations.

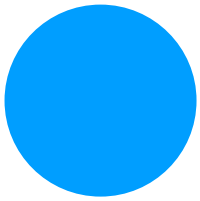
The range of data analyzed is extremely broad. It includes - but is not limited to - investors' positions, balances, investment performance information, demographics, competitive holdings, and client portal activity. This data can help identify patterns and trends that investors who have left one wealth company for another have in common. Using this model, the program has created a 95% reliable loyalty indicator that helps firms launch timely retention strategies for clients with low scores, or suggest products that have the potential to strengthen the loyalty of those with higher scores.

System benefits

The present and future uses of the suite of programs resulting from the alliance between *Fligoo* and *Broadbridges* will increase the capacity and performance of advisors, firms and their clients' investments. Together, they believe that hyper-personalization of investor needs is the key to increasing both investor satisfaction and assets, deepening and broadening a business relationship that can now be quantified between client and advisor.

Technology used

Fligoo creates software suites using advanced analytics, continuous monitoring and closed-loop feedback. With this arsenal, the alliance with *Broadbridges* uses predictive analytics mechanics by using its users' data to create a picture of their precise habits and trends.



Appendix 3. Lessons learned from deliberation with consumers

The consultation process

Although there is substantial literature on the ethical issues of AI, it does not sufficiently take into account the real expectations of citizens. The reports published by financial institutions don't reflect the sometimes divergent interests of consumers either. This is why the AMF wanted to give a voice to consumers in order to prompt the reflection on the issues of responsible AI in finance.

Three deliberative workshops, lasting 2.5 hours, were organized by the University of Montreal (Algora Lab) with the Authority on April 20, 22 and 27, 2021. During these workshops, five use cases in the form of sales pitches were presented. Fifty-one participants joined the exercise.



The objective of the consultation was to identify the most pressing expectations of consumers regarding the development of AI in financial markets. The workshops were organized in such a way as to allow time for reflection on the ethical issues of AI and time for the formulation of recommendations.

Use case scenarios

Interactive life insurance: *Bodity*

Scenario

What could be better than life insurance that knows you? Insurance that grows with you. With Dynamic Insurances' Bodity program, the first 100% interactive life insurance, you'll have peace of mind.

With us, there are no embarrassing questions about your medical history, no visits to the doctor. Thanks to artificial intelligence, Bodity gives you the choice of several state-of-the-art smart watches, connected to our application. With it, you'll be able to measure your physical activity, your movements, your sleeping hours and record your diet. You'll even get personalized advice from health experts on how to reach your goals.

Earn points with your health, and receive great rates on your life insurance, discounts and gift cards from our many partners, competitive pricing on healthy grocery items, and even free subscriptions to other services that can help improve your lifestyle. Earn points for your good habits and save on your life insurance.

Description	Setting the scene
<p>Dynamic Insurances only sell interactive plans for life insurance. Thanks to the smart devices provided with its plans, Dynamic Insurances is able to base its prices on the data captured by the devices worn by their customers. Policyholders can get discounts on their premium if they reach exercise targets that their watch records. They also have the chance to get special prizes such as gift cards from various stores if they register their exercise and health food purchases in the Bodity app.</p>	<ol style="list-style-type: none">1. You're on a tight budget and looking to reduce your monthly expenses. You're pretty active, and this type of life insurance could help you achieve that goal. Would you be willing to wear a smart watch at all times to determine your life insurance pricing?2. You are rather sedentary, you don't really like running, it even makes you anxious. Yet you could reduce the cost of your life insurance just by putting on your running shoes. Do you think it is legitimate for an insurance company to encourage you to change your lifestyle in this way?

Telematics: *Mesur-O*

Scenario

Are you an outstanding driver, known for your caution and respect for the rules behind the wheel? When your friends move, are you the one they entrust with the delicate task of driving the truck? It's time for you to benefit from your good driving skills, with the Assurance Patrimoine Mesur-O program. If you're already a client, simply install our Measure-O application on your smartphone and let its sophisticated artificial intelligence program witness your exemplary driving. You could get up to 30% off your auto insurance policy. So, if safe driving comes naturally to you, and you'd like to know that you're helping to make the roads safer, install Mesur-O now and get a 5% signup discount.

Description

Mesur-O is a telematics program that analyzes driving habits and provides personalized feedback to members. The program customizes the risk level of the car insurance according to the consumer's driving. The score is based on driving habits and the evaluation of criteria such as: speed, acceleration, hard braking and phone distraction. Mesur-O also takes into account the distance driven and the time of day. After 100 days and 1000 kilometres of analysis, the client receives a personalized premium that is reduced or increased according to the driving.

Setting the scene

1. You are often on the road, and when you are alone in your car and the coast is clear, you sometimes get reckless and speed. Since you are not putting anyone in danger, you consider yourself a responsible driver. Would you be willing to give up these moments of escapism to conform your driving to the recommendations of Mesur-O, thereby obtaining a healthy discount?
2. You join the Mesur-O program because you know you drive safely and can easily get a good discount. However, you live in a large city and sometimes other people's aggressive driving forces you to adjust your driving, which negatively affects your score. Do you feel that the criteria that Mesur-O evaluates are legitimate and representative? Would you rethink signing up for Mesur-O?

Cybersecurity: Cerberus-247

Scenario

Canadian Family Bank (CFB) is proud to be able to guarantee its clients' peace of mind thanks to its state-of-the-art security measures. With the Cerberus-247 artificial intelligence system, it's as if your account and personal information have their own bodyguard day and night. Cerberus-247 knows you and learns to identify your banking habits by analyzing the frequency of your purchases, the type of transactions you make and your interests on social networks. At the slightest sign of unusual behaviour, we will be able to act in your best interest to ensure the safety of your assets and protect your personal data. At CFB, your peace of mind is our treasure.

Description

Cerberus-247 uses cognitive security, an AI application based on human thought processes, to recognize threats. Through the interactions that trained it, Cerberus-247 proactively teaches the model to detect and analyze hazards in order to provide guidance to analysts who can perform the necessary defensive actions based on the information provided by the AI. This allows analysts to respond more quickly, eliminating time-consuming research and analysis tasks from their workload.

Setting the scene

1. Your financial institution has just suffered a massive hack of its databases. Millions of customers have had their personal information compromised, including yours. You're thinking of switching to another institution. Would you be tempted to go to CFB, since it uses the Cerberus-247 system?
2. You make regular Internet transactions with your CFB credit card. You have an embarrassing hobby and you know that Cerberus-247 learns from your transaction history to operate. Does CFB's use of Cerberus-247 make you feel uncomfortable or more confident?

Investment assistance: *InvestIA*

Scenario

You want to invest in green innovation companies and you are ready or willing to take risks, but not just any risks! The InvestIA app helps you find gems and more. Not only does InvestIA predict the performance of companies with its advanced algorithms and select the most profitable investments, it can provide arguments for the best investment according to your criteria and investment policy. Five investment leaders have already adopted it and promoted it to voting members of their Board of Directors. Transparent, fair and interpretable, InvestIA is your greatest asset to optimize your investments and those of your clients.

Description

InvestIA's goal is to predict which companies and technologies in development are most likely to offer an attractive return on investment. InvestIA evaluates about 50 parameters such as stock price, clinical trials, intellectual property ownership, or research funding. In the test phase, InvestIA has proven to be a useful tool to detect early warning signals that a company is in a poor position to make a risky investment. InvestIA's analysis allows investment companies to have a strategic investment advantage to position themselves as a leader in a specific economic sector (e.g., biotechnology). InvestIA is able to make decisions and vote on the board of directors of these companies.

Setting the scene

1. How would you react if you were a member of the board of directors of an investment firm and the CEO decided to have InvestIA sit on the board with voting rights?
2. How would you react if you learned that the investment you were expecting to develop your green solution to the office waste recycling problem had been blocked by InvestIA?

Credit scoring: PrestIA

Scenario

PrestIA, the artificial intelligence that helps you finance your projects! You have ideas, you're ready to launch your start-up, and you just need a little help, a seed fund. Unfortunately, you don't have a credit history yet or you don't qualify for a traditional bank loan. With the PrestIA app, the Canadian Family Bank (CFB) opens the doors to credit for you. PrestIA evaluates your financial situation and assigns you a credit score using the data on your smartphone: your social networks, the apps you use or the photos you take, and of course, your banking transactions. No more files to prove you're a responsible customer! Just be yourself and let PrestIA negotiate your loan.

Description

PrestIA offers both online and mobile services. Users can install the company's application on their smartphones, so even if an applicant has no credit history, PrestIA can provide a score, by analyzing their digital footprint (social media activity, geolocation and search engine activity). The software uses natural language processing to analyze users' social media posts and browser entries for indicators of liability or risk-taking. This information then informs the predictive analytics algorithm, which creates a credit scoring. Banks and credit unions can then use PrestIA's scoring to better understand the risk of insolvency.

Setting the scene

1. Your child comes to you to tell you that they have found a solution to help them get financing quickly and introduces you to PrestIA, which they've just downloaded. How would you react?
2. You are a financial services advisor at CFB. Your client asks you about PrestIA's low credit score. They ask you to review the decision because your client feels that the decision is biased. What do you do?

Summary of consumers' contributions

In this section, we group consumer reflections on ethical issues according to the AI functions identified above: evaluation, incentive, optimization. The information function does not a priori present a major issue, except for its reliability; the deliberation focused on the other three functions.

The ethical risks of evaluation

Sector: Deposit and credit

Workshops: #1 and #2

Use case: #2

Autonomy

Autonomy in the context of credit relates to the ability of consumers to give informed consent to loan conditions that they may not be able to understand accurately. Strategies for offering credit to citizens who are potentially ill-equipped in terms of financial understanding and their digital environment could push a class of people into early or inappropriate debt. Digital and financial literacy skills are necessary before taking steps with a credit application in order to make informed decisions that are adapted to their personal situation. Young consumers are particularly vulnerable to this type of app: they may be lured by seemingly attractive offers that are actually toxic and unreasonable. They could, at a young age, be caught in a spiral of over-indebtedness, which could severely delay major life plans such as starting a higher education, making a first real estate purchase and starting a family.

Privacy

The nature of the data that credit platforms and applications use to assign a credit score can raise questions about user confidentiality and privacy. Users must give up access to sensitive information such as text, call and digital transaction histories, geolocation data and address books. With regard to this data collection, consumers will have to be on the lookout for terms and conditions that they sometimes accept blindly: some companies are more explicit than others on their privacy policies and the way they use them. They will also have to assess whether they trust the lending companies that will become his creditors, whether they consider the data they will provide as relevant to the company, and make sure it is encrypted and adequately protected. The privacy issue, then, also requires significant financial and digital literacy skills development.

Well-being

The intimate nature of the data that is collected by credit applications could make the consumer feel like they are under constant evaluation, which can interfere with how they live their life as they see fit. By being monitored, the human who feels vulnerable to the judgments of a third party may feel bullied and find it difficult to feel like themselves and enjoy a life they consider worth living. The relevance of the data collected for a chance at credit may not be a fair trade for the consumer, who reveals their privacy for an opportunity to get into debt.

Lack of financial and digital literacy can also make consumers prey to potential predatory techniques of some financial institutions offering such applications. For example, a consumer who is ill-equipped to deal with over-indebtedness can have their mental health greatly affected by the stress that a precarious financial situation can generate, thus affecting their well-being.

Equity

Like many AI applications, credit scoring applications are not immune to introducing discriminatory biases into their credit analysis. For example, using users' zip codes as a demographic rating criterion, a data point often correlated with discrimination of certain predominant diasporas from certain neighbourhoods.

Also, non-traditional credit scoring applications can create access to credit for people who would not normally have access to it - or would have difficulty accessing it. On the surface, it is a product that appears to promote fairness and equal opportunity, for example, to start a business. When no love money is available for a young entrepreneur's start-up fund, this product can create an opportunity for them to achieve their goals. However, it also allows a potentially financially and digitally illiterate clientele - and therefore vulnerable to these applications - to fall into a spiral of debt, with all of its attendant problems.

Explainability

Despite the ease of use of these applications, the consumer may still be refused a loan following the evaluation of their file, and these decisions must be justified and explained. This step is necessary so that the consumer doesn't find themselves in front of a closed door following a refusal, so that they can take the necessary measures to correct their position. Although they are rarely requested, explanations as to why a file was accepted could also be transmitted, for the sake of fairness and transparency, but also to give the consumer the opportunity to know the behaviours and information that were used to their advantage in their application.

Responsibility

Even if an application is able to make, or even explain, a decision by itself, a human third party should always be part of the final decision-making process. A machine, a code or an application cannot be held responsible for discrimination or unfair treatment of consumers, and a human must be at the end of the decision chain to address these issues and correct the situations that need to be corrected.

The ethical risks of nudging

Sector: Insurance

Workshops: #1, #2, #8

Use cases: #3, #8

Well-being

Whether it's helping users take charge of their health or encouraging safe behaviour on the roads, insurers using AI have the potential to contribute to our well-being. Quantified self apps can help maintain and promote individual well-being through physical health maintenance. Telematics applications, when widely adopted, can help ensure our collective well-being by making roads a safer network or drive.

Surveillance

However, even if the machine's programmed intention is to take care of the individual and contribute to their well-being, the individual is still entitled to feel valid and justified discomfort. The idea of being quantified and monitored on a daily basis despite the coded benevolence of the program remains a form of digital intrusion that can undermine the application's initial wellness promotion intention.

Autonomy

Despite these feelings of benevolence and protection instilled by these applications, the sense of surveillance and control can lead the user to feel as though they are losing control over their lives through the normalization of behaviour by these programs. They can also lead to a fear of being forced to depend on their machines, to the point of losing the perspective of their own limits and the need to protect their privacy. Such machine addiction, through being strongly encouraged - even coerced - to use it, can affect an individual's ability to freely choose and adhere to AI-based insurance programs.

If insurance models based exclusively on smart devices continue to gain ground and the user no longer has the option to refuse the smart device model without losing their insurance, the user's autonomy is now only reduced to the autonomy they will have with regard to the use and collection of their data.

Democracy

If the insurers choose this route regardless, they will have to be very transparent about their terms and conditions and the way they aggregate and store the data of the users they are compelling in this system. The implementation of these telematics-based and self-measurement insurance systems must be based on a social ideal that can be applied to all. There should ideally not even be an economic incentive that favours some and disfavours others; rather, the goal should be to provide positive nudges to help the insured realize their moral ideals.

Equity

In their current form, insurance programs based on quantification and evaluation ensure that everyone pays their fair share. This improves access to insurance, since it no longer relies on a rough calculation of an individual's risk. However, they have the potential to reinforce inequalities. When it comes to self-measurement, an individual with pre-existing health problems, or from a more affluent background that has given them access to better education, will have less trouble complying with the application's

recommendations and accessing the rewards. When it comes to driving, an individual with exemplary driving habits, but who cannot afford a smartphone supporting the insurer's telematics applications, will not have access to the discounts that their good driving behaviour could bring. However, this precarious individual could greatly benefit from reducing one of their recurring payments.

Responsibility

When there is a malfunction or bias in the application of its applications, insurers must take responsibility for correcting the situation. They must put recourse mechanisms in place for consumers. This insurance model greatly favours the insurer, not the consumer, who must submit to constant scrutiny of their behaviours in order to reap the benefits. AI should have no other task than to evaluate and quantify. The slightest malfunction of an app can put the consumer in a very bad position with their insurer if there are no safeguards against discrimination and bias in place, and a human must immediately be in action to remedy problems.

The ethical risks of optimization

Sector: Cybersecurity

Workshops: #5

Use cases: #1, #9

Responsibility

The duty of financial institutions is to put in place all the necessary cybersecurity measures to protect consumers from the damage that hackers could do. The company should never place the burden of responsibility on the consumer.

However, the measures taken by financial institutions must not leave consumers to believe that they can take no responsibility for their personal protection and the use of their data. They must be equipped and educated to understand their digital footprint and what they need to do to ensure that their data is not compromised during their online activities.

Trust

An informed consumer will understand the division of responsibilities between themselves and the financial institutions. They must be able to trust and understand the programs they use and the way they use their data. In order to maintain this bond of trust with the machine, they must always be able to turn to a human when necessary, to counter a maneuver following an error in the program. (e.g.: closing an account following a suspicion of fraudulent activity). It must also be always possible for the consumer to personalize the system analysis and apply exclusion criteria.

Privacy

Strongly linked to trust, consumers are very concerned that the data they disseminate will be collected excessively by financial institutions under the pretext of security. They fear that it will be sold without their consent or misused in the name of security.

Sector: Investment

Workshops: #6, #7

Use cases: #5, #6

Trust

With investment AI only in a corporate context, it is difficult to ensure that it is trustworthy and remains focused on the mission of the company it supports. These AIs shouldn't be used in governance processes and mechanisms, and shouldn't be given any decision-making rights. They should only be used as a decision support tool, never as a justification. Humans should never rely entirely on the decisions of a machine, and should always be able to reason about the nature of their choices. The accountability aspect of humans should always be preserved, especially because you can't have it with AI, which you can't interact with.

Explainability

Concerns about explainability stem in part from the “black box” problem that machine learning algorithms experience. Since they are designed to program by themselves, it is currently difficult, even for the designer of an algorithm, to justify the final decisions of the machines. The black box problem causes major ethical concerns when an algorithm is part of decision-making processes that impact the life course of an individual or company. For example, if the funding of a start-up is based on the decision of a decision support algorithm and its analysis turns out to be negative, it is essential for the start-up’s owner to know why, in order to make the necessary corrections, improve and continue their entrepreneurial venture.

Explainability also goes hand in hand with transparency: by not being able to reason with the entity that refuses us, the need to debate and to have explanations about the decisions made is in vain. It is, however, necessary to satisfy this need, from the point of view of mental equilibrium, in order to properly digest a refusal. Acceptances should also be explained: a decision that can be understood can contribute to the personal realization of an individual, whether it is negative or positive. Without an explanation, it is impossible to reflect on one’s successes or failures and progress at a personal or corporate level.

Security

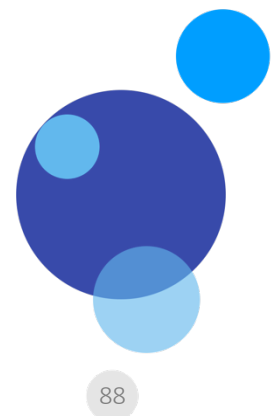
AI that evaluates projects for venture capital firms relies on a lot of data to do its job. Firms should keep the data of the companies and projects they evaluate only as long as the AI is working. The risk of it being used for personal gain (e.g., insider trading) by an individual with access to it should be zero.

Empathy

The use of AI for practically mathematical decisions is the perfect opportunity to leave emotionality and subjectivity behind. Making business decisions solely based on rationality and logic is one of the spearheads of AI, which even has a nepotistic tendency: Yuval Nova Harari notes in his book *Homo Deus* that the VITAL algorithm had this tendency, favouring in its decisions companies that gave more authority to algorithms. Machines have their own biases, and dehumanizing business and investment decisions by removing the possibility that they are tinged with empathy puts the humans who undergo the decisions at risk.

Conservatism

In the workshops, participants were shocked by the use case where AI was considered the equal of a human being. No matter how many ethical principles and how much people want to do the right thing when implementing AI in a new environment: if consumers are uncomfortable with the use of a machine or program, they will not trust the companies that use it. Digital literacy education is important to help the social acceptability of the development of AI use by financial institutions.



References

Autorité des Marchés Financiers. *Financial Glossary*. <https://lautorite.qc.ca/en/general-public/financial-glossary>

Charpentier Arthur, Michel M. Denuit et Romuald Elie. 2015. « Segmentation et Mutualisation, les deux faces d'une même pièce ». *Risques. Les Cahiers de l'assurance*, n°103 : 19-23.

CPA Canada. 2020. "Comments on the use of artificial intelligence and machine learning by market intermediaries and asset managers". Authors: Dilhac, Marc-Antoine, Rhéia Khalaf, et Manuel Morales.

Daniel, Jean-Pierre. 2016. « Et si l'assurance disparaissait ? », *Risques. Les Cahiers de l'assurance*, n° 108: 123-127.

Dilhac, Marc-Antoine, Christophe Abrassart, and Nathalie Voarino (eds.), *Report of the Montreal Declaration*, 2018 . Québec : Université de Montréal. https://www.montrealdeclaration-responsibleai.com/files/ugd/ebc3a3_5c89e007e0de440097cef36dcd69c7b0.pdf

Dillon Reisman, Jason Schultz, Kate Crawford et Meredith Whittaker. 2018. "Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability". *AI Now*. New York. <https://ainowinstitute.org/aiareport2018.pdf>

Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy, et Madhulika Srikumar. 2020. "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI", Cambridge, Harvard: *Berkman Klein Center for Internet & Society*.

Goodman B, Seth Flaxman S. 2017. « European Union Regulations on Algorithmic Decision-making and a 'Right to Explanation'". *AI Magazine*, Vol 38 (3): 1-9. <https://arxiv.org/abs/1606.08813>.

IBM. 2019. *Everyday Ethics for Artificial Intelligence*. IBM. <https://www.ibm.com/design/ai/ethics/everyday-ethics/>

Institut de la technologie pour l'humain. 2020. *Intelligence artificielle, solidarité et assurances en Europe et au Canada*. Paris. <http://www.optictechnology.org/images/files/Rapport-Complet-IA-Solidarite-et-assurances-HTF-FR-compressed.pdf>

IOSCO-OICV. 2020. *The Use of Artificial Intelligence and Machine Learning by Market Intermediaries and Asset Managers. Consultation report*. CR02/2020. <https://www.iosco.org/library/pubdocs/pdf/IOSCOPD658.pdf>

Jobin, Anna, Marcello Lenca et Effy Vayena. 2019. "The global landscape of AI ethics guidelines" *Nature Machine Intelligence*, 2019, n°9: 389-399.

Katsh, Ethan et Orna Rabinovich-Einy. 2017. *Digital Justice: Technology and the Internet of Disputes*. Oxford: Oxford University Press.

Keller, Benno. 2020. *Promoting Responsible Artificial Intelligence in Insurance*. Genève: The Geneva Association.

Landes, Xavier. 2015. « How Fair Is Actuarial Fairness? ». *Journal of Business Ethics*, Vol. 128, No. 3: 519-533.

Lundberg, Scott M., et Su-In Lee. 2017. “A Unified Approach to Interpreting Model Predictions”, *31st Conference on Neural Information Processing Systems*: 1-10. <https://arxiv.org/pdf/1705.07874.pdf>

Mackenzie, Michael, et Billy Nauman. 2021. “BlackRock pushes companies to adopt 2050 net zero emissions goal”. *Financial Times*, 26 janvier 2021. <https://www.ft.com/content/a71feaac-d3f4-4e76-a60c-c68924b06dfd>

Matthew U. Scherer. 2016. « Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, And Strategies », *Harvard Journal of Law & Technology* 29 (2): 353-400.

Microsoft. 2019. *Responsible AI Principles*. Microsoft. <https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimar6>

Monetary Authority of Singapore. 2018. *Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of AI and Data Analytics*. Singapore.

Morley, Jessica, Luciano Floridi, Libby Kinsey et Anat Elhalal. 2020. “From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices”. *Science and Engineering Ethics*, Vol. 26: 2141–2168. <https://doi.org/10.1007/s11948-019-00165-5>

Nilsson, Nils J. 2010. *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge: Cambridge University Press.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.

OECD/OCDE. 2019. *Artificial Intelligence in Society*. Paris: OECD Publishing.

OECD/OCDE. 2021. *Recommendation of the Council on Artificial Intelligence*. Paris: OECD/LEGAL/0449.

OECD/OCDE. 2022. *Framework for the Classification of AI systems*. Paris: OECD Publishing. <https://doi.org/10.1787/cb6d9eca-en>.

Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge: Harvard University Press.

Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, et al. 2016. *Artificial Intelligence and Life in 2030. "One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel*. Stanford: Stanford University. <http://ai100.stanford.edu/2016-report>

Reed Chris. 2018. « How should we regulate artificial intelligence? ». *Philosophical Transactions A of the Royal Society*: 1-12. <https://doi.org/10.1098/rsta.2017.0360>

Thaler, Richard. H. et Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven: Yale University Press.

U.S., National Science and Technology Council, Committee on Technology. 2016. Washington: *Preparing for the Future Of Artificial Intelligence*.

UK House of Lords Artificial Intelligence Committee. 2018. *AI in the UK: ready, willing and able?* [7]. Londres.

Université de Montréal. 2018. *The Montreal Declaration for the Responsible Development of Artificial Intelligence*. Québec : Université de Montréal. <https://www.montrealdeclaration-responsibleai.com>

Authors

Marc-Antoine Dilhac

Rhéia Khalaf

Nathalie de Marcellis-Warin

Manuel Morales

Robert Normand

Gabrielle Veilleux-Verreault

