afme/
Finance for Europe

Considerations on the Ethical Use of Artificial Intelligence in Capital Markets

November 2018

Association for Financial Markets in Europe
www.afme.eu

Disclaimer

Considerations on the Ethical Use of Artificial Intelligence in Capital Markets (the "Report") is intended for general information only and is not intended to be and should not be relied upon as being legal, financial, investment, tax, regulatory business or other professional advice. AFME doesn't represent or warrant that the Report is accurate, suitable or complete and none of AFME, or its respective employees shall have any liability arising from, or relating to, the use of this Report or its contents.

Your receipt of this document is subject to paragraphs 3, 4, 5, 9, 10, 11 and 13 of the Terms of Use which are applicable to AFME's website (available at http://www.afme.eu/en/about-us/terms-conditions) and, for the purposes of such Terms of Use, this document shall be considered a "Material" (regardless of whether you have received or accessed it via AFME's website or otherwise).

November 2018

Executive Summary

The AFME AI Task Force

In 2018, AFME established an Artificial Intelligence (AI) Task Force with the objectives of increasing awareness of AI in capital markets and supporting the development of future policy. These objectives are intended to enable the continued successful adoption and utilisation of this technology for the industry.

The Task Force developed its first white paper in April 2018 to: explore capital markets use-cases and benefits of AI; identify areas where AI may impact the risks faced by financial institutions; and propose control principles for managing those risks.

Considerations on the Ethical Use of AI in Capital Markets

This white paper has been developed by the Task Force to explore the ethical considerations of AI.

In our first paper we highlighted that AI capabilities are not new to financial services, and that firms have in place mature codes of conduct and controls to manage their responsibility to protect and treat clients fairly. However, the paper also identified elements unique to AI as a technology – its capacity for adaptive behaviour - which merit further ethical consideration. In this paper we explore these ethical considerations which we have grouped into two themes:

- **Data Input and Design** (e.g. AI acting on data set bias), and
- **Understanding and Control** (e.g. a lack of transparency of AI related decisions).

We have made recommendations for how firms can address the ethical considerations of these two themes. In summary firms should:

- Apply a critical view of the data sets used for each AI application and perform specific quality control checks from design through to operation;
- Train design teams and individuals to be aware of, and mitigate, biases within the function and design of individual AI applications;
- Assess what level of 'explainability' is necessary for each AI application and take this into account in its design; and
- Ensure that all AI applications are subject to a suitable control framework and audit process throughout their lifecycle.

AI is a fast-developing technology which is already being applied to specific uses within firms, as well as moving towards larger-scale deployments. Continuing discussions and information sharing on the ethical considerations of AI will be critical for the development of the capability within capital markets.

AFME looks forward to supporting the European Commission AI High Level Expert Group in developing draft AI ethics guidelines for the European AI Strategy in Q4 2018.

1. Introduction

The first AFME AI Task Force white paper in April 2018 highlighted that research into, and the adoption of AI in financial services, and specifically capital markets, has increased in the last five years. This increase has been driven largely by significant developments in computing power and the exponential growth of available data. AI technology is already being applied to many different uses, extending across the majority of functions of capital markets banks, from front office algorithms for stock selection (e.g. development of investment/structured products), to risk and compliance surveillance (e.g. identification of trading patterns that may indicate market abuse). As well as bringing benefits to banks in terms of speed and scale of processing, AI also has the potential to transform the client experience by allowing firms to tailor their services and products more precisely, or by strengthening security and resilience.

The white paper identified that, while AI is not new to capital markets, the capability and underlying technology continues to change at pace. This has the potential to impact many different market participants and, in some cases, amplify risks which have long been familiar to banks.

The development of AI is also occurring in parallel with two key trends. The first is a growing public and media attention on the technology and its potential risks and benefits. The second is increased awareness amongst individuals as to the value of their data, their rights over its use and the need to hold firms to account in this respect. Combined with the regulatory obligations of the General Data Protection Regulation (GDPR) [1], which came into force in May 2018, these trends mean that how data is used has become critical for the development of technologies such as AI.

This, in turn, has prompted a debate on ethical considerations for AI use, specifically on how to address the potential impacts of the technology on transparency, fairness, privacy and liability, whilst maintaining creativity and innovation in the development of new products. The World Economic Forum stated in 2016 that the growing emergence of AI is as much a new frontier for ethics and risk as it is for the technology itself[2]. This is further supported by work currently underway by policymakers and regulators, for example the European Commission strategy on AI published in April 2018[3], and the creation of a High-Level Expert Group in AI (AI HLEG) which proposes to draft AI ethics guidelines by Q4 2018[4].

[1] Regulation (EU) 2016/679
[2] https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/
[3] https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe
[4] https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence

2. Ethics and AI in Capital Markets

What constitutes ethics, both within the context of AI and more broadly, is complex and subjective to individuals and societies. At a high level, ethics is defined as the moral principles which individuals and organisations apply in order to make decisions[5]. The term is often used to determine the responsibilities of individuals and organisations, and what is perceived as 'right' or 'wrong' behaviour.

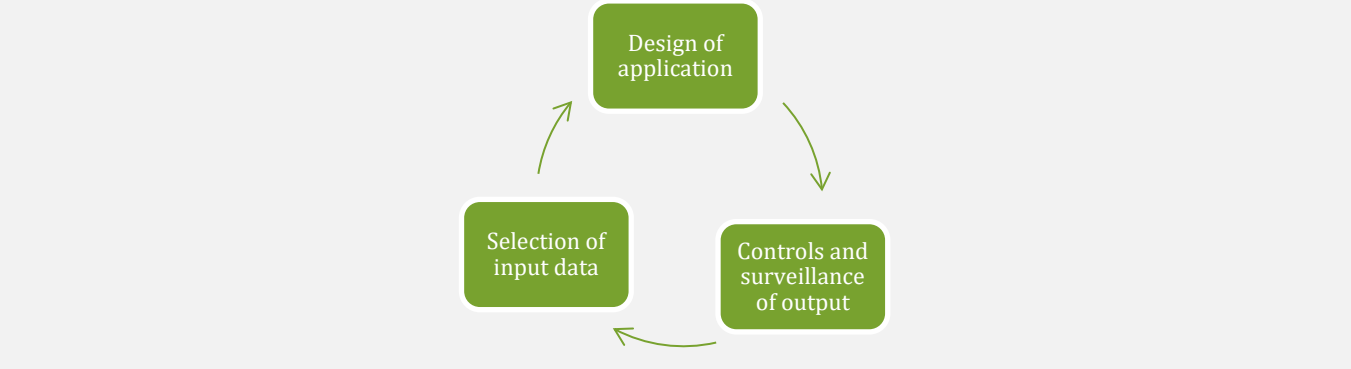
Within capital markets, the current and future use of AI should be subject to existing regulatory expectations. Additionally, in their use of the technology, firms should also be held to high conduct and ethical standards. Capital markets banks already have codes of business conduct which include ethical principles or have separate, dedicated codes of ethics. These codes outline the responsibilities and obligations on a bank's individual employees' and on the overall bank, covering areas such as: complying with applicable laws and regulations; exercising fair judgement; and executing activities openly and fairly. They are designed to address significant risks that banks face, such as systemic, customer and reputational risks, and are reviewed regularly to ensure that they keep pace with developments in technology and markets and with shifts in ethical and cultural expectations.

For AI and other algorithm-driven applications, the obligation to meet high ethical standards is not new but should continue to be embedded in any use of the technology. For example, this includes (in no particular order):

- Treating clients fairly;
- Protecting the banks and its clients from market abuse or financial crime;
- Upholding market integrity and not exposing the market to unmanageable or event systemic risks; and
- Acting as a responsible employer to upskill existing roles and ensure the right expertise for the workforce of the future.

Robust governance and risk and control frameworks are considered a requirement to ensure that all types of technology are suitably developed, deployed and monitored throughout their lifecycle. The high-level AI lifecycle can be illustrated as shown in Figure 1 below. A change to any one of the three lifecycle elements (selection of input data, design of application, controls and surveillance of output) is likely to result in, or require, changes to the other two elements.

Figure 1 – The Lifecycle of an AI application



In our first white paper we identified a series of risk categories that are relevant to consider in relation to the development and use of AI[6]. Many of these risk categories are common across other types of technology and are therefore likely to be well embedded into existing frameworks. These are therefore not covered in this paper.

However, two broad risk themes for AI were identified which we believe require further ethical consideration. This is because they are to some extent more closely aligned to the capabilities of AI given its capacity for adaptive behaviour.

The risk themes are:

- **Data Input and Design** (e.g. AI acting on conscious or unconscious dataset bias); and
- **Understanding and Control** (e.g. a lack of transparency or understanding as to why AI decisions have been made).

In support of the work underway by the AI HLG, this paper explores the ethical considerations of these two risk themes in further detail and provides recommendations on how the potential risks presented could be addressed.

[5] Oxford English Dictionary Definition of Ethics: https://en.oxforddictionaries.com/definition/ethics
[6] In our April 2018 white paper we identified examples of risks associated with AI across categories including: operational, market, financial, third-party and people

3. Data Input and Design

The availability, quality and scale of data and its use within capital markets has increased exponentially in recent years, driven partly by new regulatory requirements, such as the Markets in Financial Instruments Directive and Regulation (MiFID/R)[7], Market Abuse Regulation (MAR)[8] and the 4th Anti-money Laundering Directive (4AMLD)[9]. This, combined with the European General Data Protection Regulation (GDPR), has placed an increased focus on data management, with data becoming a significant and valuable resource that can and should be better utilised for the benefit of all market participants[10].

In our first white paper we identified that AI is reliant on large, high-quality data sets, important for:

- The initial design of any AI application, including the establishment of any necessary parameters and rules within which it must operate; and
- The ability to 'train' and test the application, and then to allow it to determine its course of action on an ongoing basis. This includes the ability of AI models to adapt their activity based on new data, which is critically important in the context of the AI lifecycle (as shown in Figure 1).

Ethical considerations for data input

The reliance of AI on large data sets creates a dependency of the AI application on the quality of the data it is given. Where that data is inaccurate, biased or not representative of a sufficient sample size, the AI application may produce results that are unfair, inaccurate or incorrect.

This is a key consideration for capital markets firms where the data they use for AI applications may pertain to clients and client activity. There may, therefore, be unintentional effects on a firm's clients, which could cut across the business principles of putting clients first and treating them fairly.

For example, if a bank trains an AI trading algorithm on data that is not representative of the full client set for which it will be used, this may result in adverse outcomes for those clients.

Recommendations:

- To mitigate the risk of AI applications making suggestions and decisions where there are issues in the data, banks should take a critical view of the data sets that are used as input for each AI application. As part of their data governance frameworks, banks should identify which data is necessary and/or relevant and perform specific quality control checks from design through to operation.
- Data sets should be representative of the wider population relevant for the use case, and do not favour particular subsets. Furthermore, banks should also input controls with the aim of preventing the AI application from replicating or introducing discrimination.
- However, while the fair treatment of clients should always be a priority, this should not be equated with any requirement to provide the same outputs for different types of clients. This would impact on the effectiveness of the AI models, whose results respond to mathematical processes on the input data. The use of representative and quality samples in the design phase will help to minimise the risk of unfair treatment. Nonetheless, where the output of an AI application is challenged, institutions should have in place suitable control mechanisms to solve any issues that may arise.

[7] Directive 2014/65/EU and Regulation (EU) 600/2014
[8] Regulation (EU) 596/2014
[9] Directive (EU) 2015/849
[10] More information is available in AFME' recent report on Technology and Innovation in Europe's Capital Markets https://www.afme.eu/globalassets/downloads/publications/afme-pwc-tech-and-innovation-in-europes-capital-markets.pdf

Ethical considerations for design

As well as being reliant on high-quality data sets, AI applications can also be limited by their design. The ethical considerations from design can manifest in a number of ways.

First, the process which the AI application is designed to replace, or augment, may have ethical considerations. Second, the team responsible for designing the AI application may, through lack of training or personal bias (whether conscious or unconscious), input or fail to mitigate potential unethical outcomes from AI applications. Finally, whether as a result of the previous point or not, the AI application may be designed with limits or parameters for its function that result in unethical activity.

For example, unless proper controls are in place, an AI trading algorithm may apply unethical trading methods to move the market in its favour.

Recommendations:

- To mitigate the risk that the design of an AI application leads to unethical outcomes, banks should train individuals in their design team to be aware of, and mitigate, biases within the function and design of individual AI applications. This should include ongoing monitoring of AI outcomes to identify any problems that may manifest over the lifecycle of the AI application.
- It is also important to be aware of the sample data that has been used for the design of the AI application, and check if a given individual/data fits the relevant population for the use case that the sample represented before applying the AI solution.

4. Understanding and Control

The nature of AI is that applications can make automated decisions about what course of action to take, using the input data they are given and the parameters they are set, and can augment processes by operating at significant speed and scale.

The term 'explainable AI' is used where the decision-making process results in specific outcomes that can be explained or described in detail by humans[11]. A range (dependent on design) exists of how far it is possible to identify and explain precisely what path an AI application has followed in making a decision. However, an explanation could be enriched with a description of the output, describing, for example, similar individuals with similar outcomes

Ethical considerations for understanding

The extent to which it is necessary to be able to explain the internal workings or decision logic of an AI application will vary depending on the function the application is performing. For example, an AI application that routes exceptions[12] to an operational process within a bank may not require a significant degree of explainable AI ('explainability'), provided that incorrect outcomes can be amended, and the application can learn from those amendments.

However, in some cases a higher degree of explainability will be necessary, for instance where the AI application has an impact on client related decisions. In these cases, it is necessary that banks (and potentially their supervisors) can monitor, evaluate and correct the output of an AI application in an appropriate timeframe. This is not only to ensure that regulatory obligations are being met, but also to ensure that the AI application is not causing unethical outcomes, for example changing the level of service a client may receive based on unknown or inappropriate factors.

For example, where an AI trading algorithm makes decisions that affect how client orders are filled, banks should be able to explain what criteria the algorithm used.

Where AI applications interact with, or directly impact on, clients raises a further ethical consideration of whether an individual should be made aware (or able to identify) that they are not interacting with a natural person, or that a human has not made a specific decision[13]. For example, where an AI application in the form of a 'chatbot' (as text or digital voice) is used by a bank to provide financial advice to a client, based on a conversation and access to their data.

Recommendations:

- To mitigate the risk an AI application may reinforce unethical behaviours, such as discrimination, firms should assess what level of explainability is necessary for each AI application and take this into account in its design. This assessment may consider:
    - The criticality of the activity being performed;
    - Compliance with regulatory obligations;
    - The interface and impact on clients;
    - Interdependencies with other internal and external systems or AI applications; and
    - The different types of actor involved in the AI application (e.g. users, consumers, etc).
- Based on the level of explainability required of an AI application and by each AI actor, different mitigation, oversight and complaints measures and mechanisms should be implemented. This could also help increase the trust clients have in their interactions with a bank's AI application.
- In addition, as required by GDPR[14], clients should have the right to ask for a decision which is not based solely on automated processing, which provides an alternative in cases where the client may not be satisfied with the explanation provided.

[11] It should be noted that even a high degree of explainability in an AI application is likely to require some level of technical expertise on the part of individuals who are responsible for monitoring or supervision, both within banks and regulators. To this end, cooperation between the public and private sectors is important to ensure that regulators and supervisors are able to keep pace with developments in banking technology.
[12] An exception is where an event, for example a trade settlement, requires intervention to complete e.g. incomplete mandatory data
[13] We note that the GDPR Article 13(2)(f) covers the obligation for data subjects to be provided with information about the use of automated decision making
[14] GDPR Article 22 (1) sets regarding automated individual decision-making, including profiling: The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.

Ethical considerations for control

AI applications are, by nature, able to adapt their results according to the input data with which they are provided. This means the initial assessments made during the design, testing and early deployment stages as to the ethical behaviour of an AI application may not remain accurate over time.

For example, even where the initial design of an AI trading algorithm includes controls against market abuse, it may not, without ongoing controls, be able to correctly identify and steer away from new variations of such abusive behaviour that emerge in the market.

An additional challenge is presented by the evolving nature of ethical and cultural expectations in relation to the use of technology. Specific uses of data, for instance, may need to be reviewed over time. Given the lifecycle of an AI application as outlined in Figure 1, the impact can be significant, particularly where such data has already contributed to the training of an AI application and therefore cannot be easily extracted.

Recommendations:

- Given the adaptive nature of AI and in conjunction with the design considerations identified, a bank's existing technology control frameworks should be adapted to monitor AI applications over their entire lifecycle, including for bias, discrimination, any decrease in explainability, or changes to what data/design elements are considered appropriate.
- Banks should consider how the AI applications are audited, including the independence of the audit team from the design team. The control framework should ensure that individuals involved in the design, approval and review of AI usage are appropriately trained and that this is continually enhanced as the technology develops. This will allow banks to establish that their use of AI continues to meet the ethical standards they have set.

Conclusion

The application of AI has the potential to transform capital markets and is already impacting many aspects of how the industry operates, from trading and client interactions to risk management and operational processing. However, AI is a rapidly evolving technology that could have far reaching impacts on society. Care must be taken to ensure its use conforms to appropriate ethical standards applied within individual banks and does not unintentionally harm the market or clients.

As part of their governance arrangements for the use of technology as a whole, banks must ensure that appropriate controls within all three lines of defence[15] can allow for the monitoring and intervention of outcomes or decisions that may reduce fairness, transparency or competition.

In this paper we have therefore considered two key risk themes and made a number of recommendations to support the ethical use of AI. In summary:

Data input and design

- Banks should apply a critical view of the data sets that are used for each AI application and perform specific quality checks, from design through to operation, to mitigate the risk of bias or inaccuracies.
- Banks should train individuals to be aware of, and mitigate, biases within the function and design of individual AI applications.

Understanding and control

- Banks should assess what level of explainability is necessary for each AI application and take this into account in its design and oversight.
- Banks should ensure that all AI applications are subject to a suitable control framework and audit process throughout their lifecycle.

It is encouraging that discussion of ethical considerations is occurring in both the public and private sector simultaneously. Continuing such discussions and information sharing in this area will be critical for the development of AI within capital markets. In particular, AFME looks forward to supporting the European Commission AI HLG in developing draft AI ethics guidelines for the European AI Strategy in Q4 2018.

[15] Three lines of defence is a risk management and controls model where the first line is management control, the second line is risk control and compliance functions established by management, and the third line is independent assurance.

Annex 1: Glossary of Terms

Glossary of common terms related to artificial intelligence (AI)	
Artificial intelligence (AI)	**The theory and development of computer systems able to perform tasks that traditionally have required human intelligence.** AI is a broad term that incorporates all terms listed below.
Algorithms	A set of rules that allow a computer to perform activities or processes to get insights from input data or to solve problems.
Black Box	A system where the internal workings are unknown or cannot be determined *(for example, a platform used to decide on a client's credit worthiness where only the broad data inputs and final decision are visible).*
White Box	A system where the decision logic is understood *(for example, a platform used to decide on a client's credit worthiness where the data inputs, decision logics and final decision are visible).*
Deep Learning	A form of neural network which is structured into a large number of processing units (normally arranged into layers).
Explainable Artificial Intelligence (XAI)	A model for AI processing, through which the decision-making process that results in specific outcomes can be explained and described in detail.
Machine Learning (ML)	An application of Artificial Intelligence that provides systems with the ability to automatically learn and improve from experience without being explicitly programmed. In short, it is a set of algorithms that allow machines to learn from data. Machine learning is typically sub-divided into three categories: supervised, unsupervised, and semi-supervised learning.
Neural Networks	"A neural network is a ML system that consists of simple interconnected processing units that are loosely modelled on neurones in the brain" *(for example, an image recognition system that learns to identify a type of image by associating certain features over time).*
Predictive Analytics	The use of current and historical data to make future predictions.
Semi-Supervised Learning	A form of 'machine learning', where an algorithm is trained on unlabelled data but receives feedback on actions taken *(for example, a surveillance platform trained on a set of transactions where some are identified as fraudulent, and a human tells the system when is has identified one correctly or incorrectly).*
Semantic Search	A system which seeks to understands the intent of search activity to improve the relevance of results *(for example, a search engine that returns results on all interest rate derivatives when asked about interest rate swaps).*
Supervised Learning	A form of 'machine learning' where an algorithm is trained on labelled data *(for example, a surveillance platform trained on a set of transaction data where areas of fraud are identified).*
Unsupervised Learning	A form of 'machine learning' where an algorithm is trained on unlabelled data *(for example, a system that learns to detect anomalies in some data without having those anomalies labelled as such).*

Notes

Bibliography

Baker McKenzie, 2017, Ghosts in the machine: Artificial intelligence, risks and regulation in financial markets

Financial Stability Board, 2017, Artificial Intelligence and machine learning in financial services

Goodfellow, I, 2016, Deep Learning: Adaptive computation and machine learning series

Microsoft, 2018, The future computed: Artificial Intelligence and its role in society

Personal Data Protection Commission Singapore, 2018, Discussion paper on AI and personal data – fostering responsible development and adoption of AI

Villani, C, 2018, For a meaningful Artificial Intelligence: Towards a French and European strategy

White & Case, 2017, Three big questions about AI in financial services

Contributors to this Paper

We are grateful to our member firms and the individuals who contributed their time and thoughts in producing this report.

AFME Technology and Operations

AFME's Technology and Operations Division brings together senior technology and operations leaders to influence and respond to current pan-European market drivers and policy.

The Considerations on the Ethical Use of Artificial Intelligence in capital markets white paper was led by the AFME AI Task Force as an initiative within the broader Technology and Operations Division.

AFME Contacts

Andrew Harvey
andrew.harvey@afme.eu
+44 (0)20 3828 2694

David Ostojitsch
david.ostojitsch@afme.eu
+44 (0)20 3828 2761

Fiona Willis
fiona.willis@afme.eu
+44 (0)20 3828 2739

/ About AFME

The Association for Financial Markets in Europe (AFME) is the voice of all Europe's wholesale financial markets, providing expertise across a broad range of regulatory and capital markets issues.

We represent the leading global and European banks and other significant capital market players.

We advocate for deep and integrated European capital markets which serve the needs of companies and investors, supporting economic growth and benefiting society.

We aim to act as a bridge between market participants and policy makers across Europe, drawing on our strong and long-standing relationships, our technical knowledge and fact-based work.

Focus
on a wide range of market, business and prudential issues

Expertise
deep policy and technical skills

Strong relationships
with European and global policy makers

Breadth
broad global and European membership

Pan-European
organisation and perspective

Global reach
via the Global Financial Markets Association (GFMA)

afme/
Finance for Europe

London Office
39th Floor
25 Canada Square
London, E14 5LQ
United Kingdom
+44 (0)20 3828 2700

Brussels Office
Rue de la Loi, 82
1040 Brussels
Belgium
+32 (0)2 788 3971

Frankfurt Office
Skyper Villa
Taunusanlage 1
60329 Frankfurt am Main
Germany
+49 (0)69 5050 60590

Press enquiries
Rebecca Hansford
Head of Media Relations
rebecca.hansford@afme.eu
+44 (0)20 3828 2693

Membership
Elena Travaglini
Head of Membership
elena.travaglini@afme.eu
+44 (0)20 3828 2733

Follow AFME on Twitter
@AFME_EU