UHASSELT

KNOWLEDGE IN ACTION

Faculty of Business Economics

Master of Management

Master's thesis

Ethical Issues in Artificial Intelligence Applications in the Financial Industry - Solutions from Auditing Services

Busra Sakinci

Thesis presented in fulfillment of the requirements for the degree of Master of Management, specialization Strategy and Innovation Management

SUPERVISOR :

Prof. dr. dr. Frank LAMBRECHTS

MENTOR :

Mevrouw Nguyen Phuong Anh DUONG

2019
2020

UHASSELT

KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

UHASSELT

KNOWLEDGE IN ACTION

Faculty of Business Economics

Master of Management

Master's thesis

Ethical Issues in Artificial Intelligence Applications in the Financial Industry - Solutions from Auditing Services

Busra Sakinci

Thesis presented in fulfillment of the requirements for the degree of Master of Management, specialization Strategy and Innovation Management

SUPERVISOR :

Prof. dr. dr. Frank LAMBRECHTS

MENTOR :

Mevrouw Nguyen Phuong Anh DUONG

This master thesis was written during the COVID-19 crisis in 2020. This global health crisis might have had an impact on the (writing) process, the research activities and the research results that are at the basis of this thesis.

Acknowledgements

Foremost, I cannot express enough thanks to my research supervisor Prof. Frank Lambrechts, and my mentor Nguyen Phuong Anh Duong for giving me the opportunity to do this research, providing helpful and valuable guidance from the beginning to the end, their immense knowledge, patience, and motivation. Thanks to them, I did not lose my motivation, even one day. I could not have imagined having a better supervisor and mentor for my thesis. Besides them, I would like to thank the rest of my thesis committee in advance for their evaluation of my work.

I would also like to thank the experts who were involved in the interviews for this research. Without their passionate participation and input, this research could not have been successfully conducted.

I would like to express my deep gratitude to Prof. Mahmut Arslan for his guidance throughout my journey on ethics and European Business Ethics Network (EBEN) Turkey and for introducing me to the field of ethics. Since my bachelor degree, Prof. Arslan has encouraged me and has given crucial support for my career steps. My sincere thanks also goes to EBEN Belgium for their support on my thesis. Our discussions during the meetings helped me a lot and increased my inspiration towards AI & ethics.

Furthermore, I consider my friends as a family. I would like to thank all my best friends from Turkey for their motivation and our good memories. Thanks a lot for your presence and all the encouragements, you are really special to me. I also want to express my gratitude to my friends that I met and had good memories during my life in Belgium.

Finally, my biggest thanks goes to my family: my parents for their unfailing support throughout my life with their love, prays, and understanding. I love you and I owe you a lot for becoming the person I am today. My dearest sister, it has been a nice pleasure to write our theses at the same time, thanks to you for your smile, your encouragement, and your guidance during my breaks and every single time. I would like to express my sincere gratitude to my family-in-law as well for their continuous support all the time. Last but not least, the completion of my thesis could not have been accomplished without the support of my husband Can. My heartfelt thanks to him for his love, support, encouragement, motivation, understanding, and caring. I cannot explain my feelings to you, thank you again for everything you did for me. P.S.: To my whole family: I feel myself so lucky for having you and I owe my success to all of you.

Büşra Sakıncı
Genk, Belgium
June 5, 2020

Executive Summary

A series of recent studies indicated that AI has been applied in financial services to save operational costs notwithstanding possible debates. Even though there are several advantages of AI, the unintended ethical consequences are inevitable. This thesis deals with the ethical consequences of AI in the financial industry with the consideration of the role of audit services. Therefore, the purpose of this research is to propose new solutions to AI and ethics in banking and insurance industry with auditing companies by proposing answers to the following questions: (i) "What is the relationship of ethical issues with AI applications in financial services (e.g. banks and insurance companies) with auditing services?", (ii) "How companies in the financial industry can solve their ethical issues, related to artificial intelligence, together with auditing services?", and (iii) "What could be improved in AI legislations in the European Union (EU)?".

As the research methodology; the grounded, interpretive qualitative study is followed together with the theoretical saturation theory after the literature is studied. During the research, nine face-to-face and remote interviews are conducted with the different experts in AI and ethics field from the financial industry (banks, insurance companies, financial services companies, and an auditing firm). The participants are considered as anonymous. All interviews are coded, categorized, grouped and analysed in order.

During the interviews, to decrease unintended ethical consequences of AI, some of the responses have supported the literature review. However, additional suggestions on unintended ethical consequences of AI have been discovered such as creating a library of AI solutions for readability, defining what unethical and ethical means to companies, considering the difficulty working with data scientists because of the different views during the collaboration with other departments, approaching AI as a solution rather than a problem, keeping a balance between explainability, transparency, and interpretability, making a honest comparison between a human being and AI system, and reversing the outcomes.

In the literature, there are specific suggestions or frameworks for internal auditors in AI applications such as the SMACTR framework, however, not specifically for external auditors. Considering the relationship of ethical issues with AI applications in financial services with auditing services, two approaches are observed: controversial and pro-active. The controversial approach corresponds to the doubt for external audits due to the fact that there are no specific standards and rules for them. On the other hand, the pro-active approach demonstrates that external audits can prevent the unintended ethical consequences of AI in the early stages because of their analytical power. They can approach more objectively and independently as the nature of their work. If auditing services are made mandatory in AI applications regarding ethical issues, the external auditors can act as certifiers by giving a guarantee label on AI applications.

As a solution for ethical issues with AI applications in financial services with auditing companies, the “Be A Hosstt” framework, a principle-based framework, is proposed in this thesis. By acknowledging the name of the framework, being a host, external auditors can act as hosts and moderators during their AI and ethics audits. Therefore, they can discover the ethical issues by controlling them. This framework can be useful for external auditors since no specific frameworks for them in the field of AI and ethics are found during the literature review. With this framework, external auditors can act as moderators and an independent certifier of AI applications to mitigate the ethical risks. This framework consists of eight elements which are consistent with each other. These elements are: “Be fair, Appreciate, Having an understanding in multidisciplinary teams,

Own challenger models, See what is hidden, Standards, Tailor-made audits, and Test". First of all, the external auditors should not forget being fair, independent and objective, after that, they should understand the purpose of the models, audits, and AI. In external audits, even though multidisciplinary teams are significant as indicated in the literature, having an understanding in these teams is also crucial to improve the results by appreciating what other team members mean. Additionally, even though there are lots of frameworks and models in the industry, as external auditors, they should have their own challenger models to be able to detect biases in the financial services companies' models. Their own challenger models should be designed considering the standards which need to be put by the European Commission (EC) into the EU legislations. Therefore, they can act as certifiers. Afterwards, during their audits, they should try to see what is hidden to protect themselves from blindness. Every company has different values and policies, therefore, external auditors should conduct tailor-made audits by adjusting their methods considering the different values of different companies. Thus, they will be able to get efficient results with their own challenger models and considering the standards, and corporate policies and values of the financial companies. As a final step, they should test their own challenger models. If the tools and the model are not working well, they cannot detect the risks properly. These elements are compatible with each other and thus, external auditors should consider them by thinking the whole process.

During the interviews, some suggestions have been discovered for the improvements of AI legislations in the EU. The EC should approach this issue more broadly, define what no-risk or high-risk applications mean based on specific criteria and reconsider the voluntary labelling system and its circumstances. Following, the EC can also reinforce the control mechanisms by specifying the necessary standards which can be considered by auditors to give a certification to banks or insurance companies, while remembering not to kill the innovation when setting the standards.

The main contributions of the thesis are the investigation of the relationship of AI and ethics with auditing companies, proposing the "Be A Hosstt" framework as a solution for ethical issues of AI applications in financial services with auditing companies, and giving additional suggestions for the needed improvements of AI legislation in the EU. In addition, the proposed framework can also be considered by the experts as a clarification of the application of the EU legislations on AI and ethics since the elements of the "Be A Hosstt" framework are compared separately for each element with the "EU Guideline as a framework for Trustworthy AI".

To suggest more solutions, future research can investigate the relationship between ethical consequences of AI applications and auditing services more thoroughly. Due to its limitations, further research can be considered by conducting more interviews to extend its relation. Last but not least, the technical mindset can be added to the "Be A Hosstt" framework together by testing it. Hence, the effectiveness of this framework can be observed.

Abstract

The purpose of this thesis is to discuss ethical consequences that can occur with the use of AI technology in the financial industry. The thesis employs a number of ethical issues related to AI in the financial industry, specifically the effects and approach of auditing services in this field. Effects of AI on customers and operations are discussed during the digital transformational era. The current and suggested solutions to decrease unintended ethical consequences of AI are introduced. The roles of internal and external auditing services, and existing regulations in AI and ethics are indicated. Following a thorough literature review, a qualitative study is carried out by conducting interviews with experts –mainly from the financial industry– in the field of AI and ethics. The responses are transcripted, and the common ideas expressed by the participants are grouped together.

By maintaining an interpretive grounded theory approach, it is concluded that the relationship of AI and ethics with auditing companies can be strengthened, if the necessary standards are released by the European Commission (EC) and the auditing of AI applications is made mandatory. Secondly, a principal-based framework for auditing companies is suggested in this thesis to help auditors. It consists of eight basic elements which are created based on the responses received during the interviews. Thirdly, even though the EC has released some principles and a new white paper on AI, it is proposed that they need to go beyond these recommendations since companies need to get specific standards in order to follow them properly. In line with the outcomes, suggestions to the European Union legislations are proposed.

Keywords: artificial intelligence, ethics, AI, auditing services, external audit, auditing framework, AI legislation, Trustworthy AI.

Author: Büşra Sakıncı
Student ID: 1953229
E-Mail: busra.sakinci@student.uhasselt.be

Supervisor: Prof. dr. Frank Lambrechts
Mentor: Nguyen Phuong Anh Duong

Contents

1 Introduction **1**
- 1.1 Background . . . 1
- 1.2 Scope of the Thesis . . . 2
  - 1.2.1 Research Methodology . . . 2
- 1.3 Contributions of the Thesis . . . 3
- 1.4 Outline . . . 3

2 Literature Review **5**
- 2.1 AI uses in the Industrial Revolution . . . 5
- 2.2 Ethical Repercussions of AI . . . 6
- 2.3 Current and Suggested Solutions . . . 7
- 2.4 The Role of Auditing Companies . . . 9
- 2.5 Existing Regulations . . . 11
  - 2.5.1 Ethics Guidelines for Trustworthy AI . . . 11
  - 2.5.2 A New White Paper on AI . . . 13
- 2.6 Conclusion . . . 14

3 Methodology **15**
- 3.1 Interviews . . . 16
  - 3.1.1 Interview Structure . . . 16
  - 3.1.2 Data Collection & Analysis . . . 16
- 3.2 Findings . . . 17
  - 3.2.1 Decreasing Unintended Ethical Consequences of AI . . . 17
  - 3.2.2 Auditing Services with Ethical Issues in AI Applications . . . 19
  - 3.2.3 Solutions to Ethical Issues on AI together with Auditing Services . . . 21
  - 3.2.4 Suggestions for AI Legislations in the EU . . . 24
- 3.3 Conclusion . . . 28

4 Discussion **29**

5 Conclusion **33**
- 5.1 Conclusions . . . 33
- 5.2 Implications for Future Research . . . 34

Appendices **35**

A Interview Questions **37**

B Categorisation of Participant Responses **41**

References **60**

List of Figures

2.1 The SMACTR framework for an internal audit. 9
2.2 Traditional auditing methodology vs. CA methodology. 10
2.3 Framework proposed by the EC. 12

3.1 Be A Hosstt Framework for auditing companies. 22

List of Tables

2.1 Literature review findings summary. 14

3.1 Interview details. 15
3.2 Second-order categories related to the unintended ethical consequences of AI. . . . 17
3.3 Second-order categories on AI & ethics with auditing services and AI in EU legislations. 20
3.4 Third-order categories on AI & ethics with auditing services and AI in EU legislations. 21

A.1 Research questions covered in this thesis. 37
A.2 Introductory questions. 37
A.3 Questions related to Research Question I. 38
A.4 Questions related to Research Question II. 39
A.5 Questions related to Research Question III. 40

B.1 Quotes related to second-order categories on unintended ethical consequences of AI. 42
B.2 Quotes related to second-order categories on the relationship between auditing services and ethical issues in AI applications. 50
B.3 Quotes related to second-order categories on the solutions of ethical issues in AI applications together with auditing services. 52
B.4 Quotes related to second-order categories on AI legislations in EU. 54

Nomenclature

AI	Artificial Intelligence
AML	Anti Money Laundering
B2C	Business to Customer
CA	Continuous Auditing
CFT	Combating the Financing of Terrorism
EC	European Commission
EU	European Union
FAT	Fairness, Accountability, Transparency
FSB	Financial Stability Board
HITL	Human in - The - Loop
KBS	Knowledge-Based System
NLP	Natural Language Processing
SME	Small-Medium Enterprises
XAI	Explainable AI

Chapter 1

Introduction

1.1 Background

People name themselves *Homo sapiens* –man the wise– due to the significancy of human intelligence (Russell & Norvig, 2014). The field of artificial intelligence (AI) works not merely to understand how humans think, but also to build intelligent entities (Russell & Norvig, 2014). According to Russell and Norvig, "If we are going to say that a given program thinks like a human, we must have some way of determining how humans think." (Russell & Norvig, 2014, p.3). At the same time, it also should not be forgotten that morality is the nature of humanity, even though there could be some human beings without moral problems. If we think about morality and ethical problems when we discuss human behaviour, then AI applications, which take the human brain as an example and try to build intelligent entities, will also experience moral and ethical problems.

European Commission (EC) has indicated the definition of AI with several aspects as follows: "It clarifies certain aspects of AI as a scientific discipline and as a technology, with the aim to avoid misunderstanding, to achieve a shared common knowledge of AI that can be fruitfully used also by non - AI experts, and to provide useful details that can be used in the discussion on both the AI ethics guidelines and the AI policies recommendations." (European Commission, 2019). Throughout the years, there has been a perplexity, curiosity and big considerateness towards AI. One of the studies from 1988 has described AI as knowledge based/expert systems (KBS), natural language analysis, and search, inference and planning procedures (Pau, 1987). From economists to banks, financial services as well as management departments, regard the stream of the use of AI technology by considering its far-reaching consequences (Pau, 1987). As AI still goes further because of the developments in science and engineering (Russell & Norvig, 2014), there are many different descriptions and approaches regarding AI in the literature. For instance, according to Russell and Norvig (2014), four approaches of AI have been particularly indicated, namely the cognitive modelling approach by thinking humanly, the Turing Test approach by acting humanly with the capabilities of natural language processing, knowledge presentation, automated reasoning, and machine learning, the "laws of thought" approach by thinking rationally and the rational agent approach by acting rationally. Different people have approached to define AI with different goals in mind. It is thus unclear which approach companies should take while defining AI, as no consensus has been reached in the literature.

Considering the descriptions above, value can be created if the company can create entry barriers or mitigate the risk of doing business (Fruhan, 1979). To achieve these goals, AI technology and experts systems can be used in the financial industry (Vasarhelyi & O'Leary, 2000). Many financial and insurance companies are using AI to increase the quality, efficiency, and competitive leverage of their operations (Rauch-Hindin, 1988). Even in such an old study, the importance of dependency on AI in the financial industry has been highlighted as follows: "Many market analysts believe that in the next decade or two, the financial and insurance industries are going to become very dependent on AI." (Rauch-Hindin, 1988, p. 300). Moreover, in the financial services industry,

for instance, banks and insurance companies use AI as a mechanism to ease their operations by reducing their operational costs such as using chatbots. Thus, financial services companies do not need physical advisors to serve their customers.

Even though there are advantages of using AI, the disadvantages far outweigh them and the financial services companies should consider those drawbacks. Within the financial services industry, there are lots of implications of AI; namely algorithmic trading, back testing, fraud detection, robo-advisory services, chatbots etc. According to Rauch-Hindin (1988), cost, training, risk, portability, hardware, security, user interfaces, and integration with traditional hardware and software will be major concerns in the financial and insurance industries due to their high-level of dependence on AI. Additionally, the digital transformation brings the complexity to the financial industry; therefore, tech giants developed bias detectors such as Google's WhatIf and IBM's bias detector to solve the problem. Furthermore, in order to establish AI legislations and create a framework to solve the aforementioned issues, The EC announced "Ethics Guidelines for Trustworthy AI" in April 2019 following "A New White Paper on AI" in February 2020. They encouraged that uses of AI should be lawful and respectful of ethical principles and should consider the social environment.

The AI software are becoming more complex day by day, due to more available data in the manner of chicken-and-egg dilemma: Are AI programs evolving just because of the data, or is more data available because of AI programs? This might be blurring; however, the financial services should take into account that AI programs certainly aid them to gain a competitive advantage, and, at the same time, bring ethical challenges.

1.2 Scope of the Thesis

There are several advantages of AI along with its considerable disadvantages. As reported by Buchanan: "Along with big data, AI is viewed in the financial services sector as a technique that has the potential to deliver huge analytical power. Yet many risks still need to be addressed. Many AI techniques remain untested in financial crisis scenarios." (Buchanan, 2019, p. 29). There are few studies in the literature about the relationships of AI and ethics with external auditing companies or related frameworks for them. When internal auditors coordinate with external auditors in assurance-related activities, a lower risk has been observed (Wang & Fargher, 2017). It is also suggested that effective coordination between internal and external auditors can positively affect mechanisms in the organisation to decrease the risks (Wang & Fargher, 2017). Therefore, this thesis deals with the ethical consequences of AI in the financial industry with the consideration of the role of audit services. More specifically, it aims to provide new insights about creating solutions with auditing companies in the field of AI and ethics in banking and insurance industry. Henceforth, the main research questions that this thesis answers are:

- What is the relationship of ethical issues with AI applications in financial services (e.g. banks and insurance companies) with auditing services?
- How companies in the financial industry can solve their ethical issues, related to artificial intelligence, together with auditing services?
- What could be improved in AI legislations in the European Union (EU)?

1.2.1 Research Methodology

Interviews are conducted with several experts in the field of AI and ethics from the financial industry in order to provide consistent results that are relevant for the analysis and results. A total of 9 interviews are conducted with participants from banks, insurance companies, financial services companies and an auditing firm. As requested by the participants, their names are not mentioned, and anonymous responses are used. The responses from the participants are analysed to find commonly expressed ideas, and these ideas are grouped to aid the discussion and arrive at conclusions.

1.3 Contributions of the Thesis

This thesis provides the readers with a thorough literature review, which indicates the usage of AI applications in banks and insurance companies, together with their advantages and disadvantages. Besides, the role of audits on AI applications and current legislations have been expressed in the literature review.

Many unintended ethical consequences of AI applications and some solutions for them have already been indicated in the literature. During the interviews, new solutions to these consequences were proposed by the participants. The main contribution of this thesis is the indication of the relationship between auditing companies and AI applications in the financial industry. Two approaches have been considered, namely "controversial" and "pro-active" approaches. As long as standards are put by the EC, which auditors can apply on their audits by giving a certificate to the companies, the control mechanism can be reinforced. Thus, the unintended ethical consequences of AI can be mitigated. As a solution for auditing companies, in this thesis, a new principle-based framework, "Be A Hosstt", has been created. In this framework, it is supposed that auditors can act as certifiers and controller in the AI applications together with its ethical risks for banks or insurance companies. Taking the principles given in the framework would help auditors during their works on AI. Another contribution brought by this study is figuring out additional suggestions for AI legislations in the EU. Last but not least, beyond being a guideline for external audits, the "Be A Hosstt" Framework can also be regarded as an understandable guideline by the professionals during the application of EU legislations.

1.4 Outline

This thesis consists of five chapters. In Chapter 2, the findings of a thorough literature review are presented. The focus of the literature review is on the ethical effects of AI applications on customers and operations in the financial industry, the role of internal and external audits in this field and the existing regulations on AI. In Chapter 3, the research methodology adopted in this thesis is explained in detail. This chapter is followed by Chapter 4, in which the research results are discussed, together with a comparison with the existing academic literature as described in the literature review section. Finally, in Chapter 5 conclusions are drawn and implications for future research is discussed.

Chapter 2

Literature Review

In this chapter, previous studies on the ethical consequences of AI in the financial industry are presented. Specific attention is given to cases that are related to banks, insurance companies, auditing firms, and their interactions. The chapter begins with an explanation of the role of AI in the industrial revolution. This is followed by the exploration of ethical repercussions of AI, where the focus is on the unintended consequences of AI. Following the different effects of AI in the financial industry, the current and suggested solutions are indicated in Section 2.3. These are then followed by the role of audits in Section 2.4, and the existing regulations in Section 2.5.

2.1 AI uses in the Industrial Revolution

From the beginning of the third industrial revolution, as a part of the digital transformation, banks connected with their customers using computers and internet (Fourie & Bennett, 2019). In the third industrial revolution era, many applications of AI-based fraud prevention systems are clearly seen in the financial services companies such as Mastercard's Decision Intelligence (DI) technology to establish a baseline of Mastercard's customers based on their past behaviour, HSBC's Quantexa AI Software to fight money laundering and NatWest's Corporate Fraud Insights from Vocalink Analytics to detect and protect fraudulent transactions (Buchanan, 2019). In this transformational era, while using AI in their services, incumbents have encountered both ambitions and concerns due to the competitive advantages and disadvantages. Below, the main competitive advantages of AI in the financial industry are discussed.

Customer support

AI is being used in various ways in the insurance sector by supporting innovative ways to keep their customers through assisting virtually with chatbots (Riikkinen et al., 2018), and calculating insurance risk (Vassiljeva et al., 2017). In banks, platforms based on chatbots and robo-advisor algorithms aid the customer choice of investments, banking products and insurance policies across financial services companies as well (Buchanan, 2019).

Detecting fraudulent transactions

Regarding banks, AI is more beneficial than Benford's Law in detecting fraud, because the algorithms can analyse millions of data points and detect blind spots of the fraudulent transactions which humans can not realise (Buchanan, 2019).

Digital banking platforms

Banks are benefiting from digital platforms in order to manage risk and safeguard from fraud, anti-money laundering initiatives, and customer identity confirmation. Socure's ID+ Platform to

predict data science on information legitimately for know your customer (KYC) purpose could be given as an example (Gossett, 2020). Another application that utilises AI is lending and risk management to optimise risk within the financial services by using digital platforms such as ZestFinance's AI-based software to develop fairer systems (e.g. downgrading credit data), and Feedzai's managing risk platform by controlling transactions used by Citibank (Gossett, 2020).

Reduce costs

Financial Stability Board (FSB) (2017) has announced that in order to get more profit, companies have boosted their adoption toward AI technology to reduce their costs, improve productivity and keep up with their competitors' adoption of AI because of the reputation (hype). For low costs and qualified advice for consumers rather than traditional investment advice, robo-advisors have been proposed in a recent study (Rohner & Uhl, 2018). Insurance companies also use AI in their operations to decrease the cost and develop insurability (Hall, 2017).

Risk management

AI applications contribute to the economic growth by increasing the efficiency of financial services with more efficient risk management, encouraging collaboration with new economies of scope –especially in e-commerce and sharing economy industries– and stimulating investments. In e-commerce, customers are analysed depending on the transaction data added to payment and settlement activities (e.g. Who buys what, when, and where?) (Financial Stability Board, 2017).

2.2 Ethical Repercussions of AI

As discussed in the previous section, AI has found different uses in the financial services. However, in today's transformational era, the financial companies need to consider the competitive disadvantages of AI as well. In the literature, many ethical consequences of AI technology have been highlighted. For instance, FSB (2017) pointed out the rapidly evolving phase of AI and the unavailable data on its usage, and concluded that the development of this technology should be controlled. Although AI has huge potential benefits, there are consequential concerns for this technology such as potential discrimination, safety, privacy, ethics, and transparency as discussed below.

Complexity

Banks are suffering from the complexity of AI in their customers' experiences. This complexity occurs mainly in their fraud detection operations, using machine learning technology to detect frauds properly and consistently (Fourie & Bennett, 2019).

Customer dissatisfaction

One of the major drawbacks is a customer dissatisfaction about realising that they are facing a machine rather than a human, because of the incapability of the chatbot (Fourie & Bennett, 2019). Another issue for banks is the natural language processing (NLP) considering the communication way, whether it is AI or human interface (Buchanan, 2019).

Discrimination

If the financial companies do not understand their systems well, it might result in dreadful chain reactions. Google's former CIO indicated this matter as: "People should be treated fairly but until now there was no way for banks to do the right thing because they could not understand their own models well enough to know what variables, if any, cause discrimination." (Fuscaldo, 2020).

Explainability

Using AI applications in the organisations must be manageable due to their potential risks. As one of the highly-regulated industries, financial services such as insurance companies are obliged to tell the reasons of unapproved insurance applications to the regulators to prove whether there is a systemic bias or not. Explainability is a considerable issue in raising the trust in AI, which is handled by demonstrating how a model works to business stakeholders (MIT SMR Connections, 2020).

Privacy

In B2C insurance, the sources of customers are considerable in the AI-driven automation since the insurers seek to use online channels in order to enable a huge capacity of the sources of customers by gaining accesses to new uses. However, they are limited by the quantity of risk (Zarifis et al., 2019) due to data quality and privacy issues (Accenture, 2017). In a recent study, the personal privacy issue has been highlighted as follows: "Artificial intelligence is an emerging technology, but the current policy system is not perfect and the supervision mechanism is not in place. This technology inevitably brings risks such as personal privacy leakage, ..." (Li et al., 2019, p. 101).

Stereotyped decision mechanisms

Depending on the assumptions made in the data that is used while training the AI, the decision making process involving AI can result in unfavourable mistakes (Cramer et al., 2018). Hence, the importance of Fairness, Accountability and Transparency (FAT) in the data-driven and algorithm-driven systems is supported in prior studies (Larsson et al., 2019). Besides the complexity, transparency, and regulatory issues, a great deal of data brings privacy issues due to the manipulation of AI by making judgments (e.g. men tend to repay loans more than women based on the historical data of credit rating models), predicting the future and creating new angles (Fourie & Bennett, 2019). To illustrate, based on the customer's ability to pay their loan and their shopping habits, their creditworthiness is determined by using AI mechanisms.

Transparency

Transparency is an essential element for the insurers as well, since the pricing policies have to be explained to the customers instead of using "The computer said so." as an unacceptable answer (Hall, 2017).

2.3 Current and Suggested Solutions

To solve the ethical consequences of AI, there are some current practices. For example, mobile payments and EMV (Europay, Mastercard, Visa), in other words smart or chip cards, assisted the improvement of card authorisation applications (Pascual, 2015). Mobile payments and EMV also helped with the reduction of false-positive declined transactions, which causes customers' infidelity and is a big threat to the financial services industry owing to the suspected transactions. Another example is that credit agency Equifax provides perceptibility to financial services companies with AI-driven scoring tools such as in-house machine learning models, which give reasons for approved or unapproved insurance applications (MIT SMR Connections, 2020). Besides current measures, in the literature, crucial suggestions have been proposed for the ethical issues encountered in AI technology as indicated below.

Adjustments on the regulations

The existence of this technology essentially requires regulatory compliance, prudential regulations, data reporting, rules on anti-money laundering and fighting against the financing of terrorism (AML/CFT) (Financial Stability Board, 2017). These requirements are inevitable and should be

extended in a detailed way, for instance, through new data standards, new data requirements, new data reporting requirements, or other institutional changes in financial services (Financial Stability Board, 2017).

Building readability of the algorithms

The regulators may build the readability of the algorithms, and then the transparency of algorithms might be achieved (Fourie & Bennett, 2019). According to FSB, uses of AI in finance are still in an immature and breakneck phase. Thus, data unavailability is considerably high, and all analysis and advancement should be considered as fundamental.

Check the inaccuracies

For the insurers together with banks, the source of the data is an issue since data quality is also significant as much as the data volume, which can be improved by checking and cleaning the inaccuracies (Zarifis et al., 2019).

Coordination in the company

The importance of coordination in the company has been highlighted as follows: "Auditors must work together on an audit and bank loans are part of a portfolio of loans and thus, must be considered in concert with other loans. Thus, value derives from vehicles designed to ensure the coordination of such efforts." (Vasarhelyi & O'Leary, 2000, p.14).

Having diverse teams

Establishing diverse data science teams, and safeguarding the following of ethical guidelines have been indicated as fundamentals to create a trust within any organisation (MIT SMR Connections, 2020).

Hire new employees

For financial institutions, a rapid development of AI skills, and the extension of their knowledge by hiring new employees who are suitable to manage the complexity of AI have been suggested in order to integrate their decisions regarding AI (Fourie & Bennett, 2019).

Human in-the-loop

By keeping the human in-the-loop (HITL), which means human-algorithm collaborations, unintended ethical consequences of algorithms can be prevented (Green & Chen, 2019). Also, in a recent study, human as a part of the final decision has been proposed as follows: "Whenever an AI-assisted process leads to biased or unfair decisions, measures such as retraining the human reviewers may be part of the solution." (Berscheid & Roewer-Despres, 2019, p.16).

Improve the systems

To avoid the mistakes, Fourie and Bennett proposed that in the fourth industrial revolution, banks should recognise their capabilities and the consistency of their strategies of AI. That's why data relevancy is significantly important, hence, the financial companies such as banks need to develop their mechanisms to predict more coherently. Companies also need to improve their chatbots systems in order to protect misunderstanding, if not it will cause customer dissatisfaction (Fourie & Bennett, 2019). Also, to solve the customer dissatisfaction issue, banks must set high security features in order to assure private and confidential information within the chatbot services since they need to increase the confidence level of customers, especially mostly millennials (Quah & Chua, 2019).

New frameworks

In the literature, some new frameworks have been proposed to mitigate the effects of AI. However, only the most related framework regarding AI and ethics, the SMACTR framework, has been considered in this research. According to Raji et al. (2020), for the further development of an internal audit framework, the SMACTR framework –acting like the North Star to mentor the development of lifecycle– was suggested to advise practical implementations. This framework seeks to set up the combining structure of audit and engineering processes in order to strengthen the responsive development of AI systems (Raji et al., 2020). An overview of the framework is given in Figure 2.1.

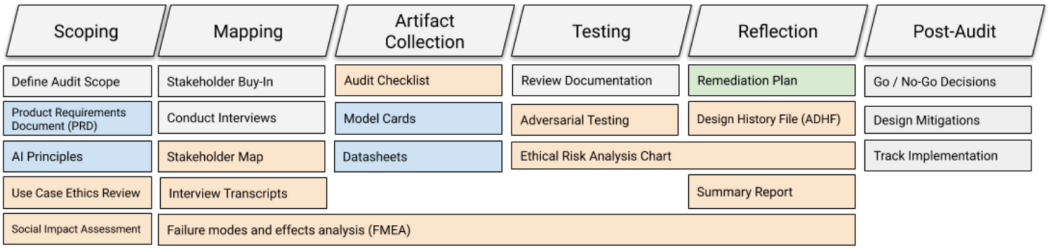

Figure 2.1: The SMACTR framework for an internal audit.

According to the SMACTR framework, there are six noticeable stages: Scoping, Mapping, Artifact Collection, Testing, Reflection and Post-Audit. The orange and blue coloured areas show the documentation processes produced by the auditor team and the engineering and product teams, respectively. The green coloured area shows the jointly developed documentation. The suggested framework is designed to add to closing the accountability gap in the development and deployment of AI systems by enclosing a vigorous process to safeguard audit integrity (Raji et al., 2020). Furthermore, as stated in the framework, to diminish the unintended social consequences (positive or negative) of AI, a social impact assessment method has been proposed in the scoping stage to inform ethical examination. This method consists of two main steps; namely an evaluation of the asperity of the risks and a labelling of the relevant social, economic, and cultural impacts and harms as the results of using AI system (Raji et al., 2020). Additionally, in the mapping stage, orchestrating with the suitable stakeholders in the audits by getting inspiration from other industries such as ethnography-inspired fieldwork is practical to increase the knowledge of the engineering system (Raji et al., 2020). Similarly, the ability to approach crucial individuals around the development process in the company by making some semi-structured interviews is fundamental in internal financial auditing in order to secure transparency (Raji et al., 2020). In the mapping stage, Failure Modes and Effects Analysis (FMEA) has also been offered as a method to identify and mitigate possible failures in various products, designs, systems and services, which is commonly applied in safety engineering (Stamatis, 2003). FMEA affects the entire risk management system by analysing the suggested work out. Besides, FMEA has not yet been practiced in the identification of ethical risks in large-scale AI models (Raji et al., 2020). In this stage, FMEA has been suggested as an analysis which should be regarded as a first concern for the next tests in AI technology.

2.4 The Role of Auditing Companies

Internal algorithmic audits are used as a mechanism to successfully implement the organisational AI principles during engineering processes (Raji et al., 2020). Before the possible negative consequences of AI technology arise, such audits can be used to obtain a desired result. Internal auditors have direct access to internal systems (e.g. intermediate models and training data, which are currently kept as companies' trade secrets), hence external audits are suffering from the lack of access to some part of the internal processes at the audited organisations (e.g. they use an

API) (Christian et al., 2014). Internal auditors –employees of the organisation– can use the audit results in the recommendations for structural organisational changes by making the whole engineering process auditable and adjusting ethical standards (Raji et al., 2020). Internal audits are consummating external accountability by developing artefacts or transparent information (Shah, 2018). Third parties such as external auditors can use this information, hence, internal audits can accomplish transparency by implementing rigorous reporting requirements (Raji et al., 2020). External audits are defined by default to detect existing social biases and risks and to perform as accountable for the models (Raji et al., 2020).

In today's innovative world, the continuous auditing (CA) methodology has been suggested in the auditing system as an additional successor to the traditional auditing methodology considering technology and automation (Chan & Vasarhelyi, 2011). As illustrated in Figure 2.2 (Chan & Vasarhelyi, 2011), the traditional auditing methodology is applied periodically and manually together with independent roles of the internal and external auditors. Therefore, CA escalates the efficiency and effectiveness of the audit process in order to get real-time assurance. CA has been proposed for the auditing of financial information for the companies. In the CA system, the internal auditor can investigate irregularities or exceptions from the system with the professional and skeptical approach. The implementation of the system can be done by internal auditors, while the external auditors can act as an independent certification provider of the internal audit CA system. External auditors can certify the audit report of the CA system if there are no detected abnormalities in the black box log. Therefore, external auditors can provide valuable counselling comments on the CA system since they obtain the knowledge and experience from the industry (Chan & Vasarhelyi, 2011).

Traditional auditing	Continuous auditing
1. Frequency: • Periodic	1. Frequency: • Continuous or more frequent
2. Approach: • Reactive	2. Approach: • Proactive
3. Procedures: • Manual	3. Procedures: • Automated
4. Work and role of auditors • Bulk of the work performed is centered around labor and time intensive audit procedures • Independent roles of the internal and external auditor	4. Work and role of auditors • Bulk of the work performed is centered around handling exceptions and audit procedures requiring human judgement • External auditor role becomes the certifier of the continuous auditing system
5. Nature, riming, and extent: • Testing consists of analytical review procedures and substantive details testing (nature) • Controls testing and detailed testing occur independently (timing) • Sampling in testing (extent)	5. Nature, riming, and extent: • Testing consist of continuous controls monitoring and continuous data assurance (nature) • Controls monitoring and detailed testing occur simultaneously (timing) • Whole popoulation is considered in testing (extent)
6. Testing: • Humans perform testing	6. Testing: • Data modeling and data analytics are used for monitoring and testing
7. Reporting • Periodic	7. Reporting: • Continuous or more frequent

Figure 2.2: Traditional auditing methodology vs. CA methodology.

Researchers have proposed that for the best gain from the auditing process, the reliability of both internal and external audits is required (Stanwick & Stanwick, 2001). The coordination between the internal and external auditors can help diminishing biases on internal auditors' fraud risk management, since the external auditors' techniques and methods facilitate internal auditors to advance their assurance-related work (IIA, 2013). To mitigate the risks in the internal audits, working with external audits has been suggested in the literature as: "We find that the internal auditors' assessed fraud risk is lower when there is a coordination with external auditors, compared

to when such coordination does not exist." (Wang & Fargher, 2017, p. 1198). Also, the regulations in the financial industry propose that internal audits are only one of the many important elements to consider. Independent outside assessors moderate the external audits in terms of the quality and other management system practices of business (Simon et al., 2014). For instance, it has been observed that the external audit reports discovered by the auditors including a great deal of comments have an analytical power on the prediction of the bankruptcy situation of firms (Muñoz-Izquierdo et al., 2019). The need for external audits has been supported as follows: "Errors, fraud and corruption could be prevented by implementing principles from the Code of Ethics and auditing (external) procedures." (Ionescu, 2017, p. 33). Hence, keeping independent and objective aspects in the audit process is needed (Raji et al., 2020).

2.5 Existing Regulations

In this section, existing regulations and recommendations for AI applications are presented. AI encompasses a black-box effect, complexity, unpredictability, moderately autonomous behaviour and a wide variety of risks such as bias and discrimination, notwithstanding its potential benefits. To diminish the risks and obscurity associated with the widespread use of AI, there is a need to create a regulatory framework in EU. To protect and control the private data, General Data Protection Regulation (GDPR) was recently constructed in the EU. To mitigate the ethical risks of AI, there is no specific regulation as in the GDPR.

According to the EC, an understandable European regulatory framework would strengthen trust in connection with consumers and businesses in AI, and hence it would endorse Europe's innovation capacity and competitiveness in this area by accelerating the adoption of the technology (European Commission, 2020). In this section, the approach of the EC and its applications are discussed.

2.5.1 Ethics Guidelines for Trustworthy AI

On 8 April 2019, the High-Level Expert Group on AI introduced Ethics Guidelines for Trustworthy Artificial Intelligence. The first draft of the guidelines was published in December 2018, bringing more than 500 comments in an open consultation (The High-Level Expert Group on AI (AI HLEG), 2019a). As stated in the Guidelines, Trustworthy AI should include lawfulness with the concerns of all applicable laws and regulations, ethics with the concerns of ethical principles and values, and robustness with the concerns of a technical perspective and a social environment. These components are conforming with the core tenets of the EU such as fundamental rights, democracy and the rule of law. There are two elements of Trustworthy AI, (i) an ethical purpose and (ii) technical robustness and reliability (The High-Level Expert Group on AI (AI HLEG), 2019a). The EC listed four ethical principles based on fundamental rights in order to safeguard the development of AI in a trustworthy manner. These principles are (The High-Level Expert Group on AI (AI HLEG), 2019a):

i) Respect for human autonomy;

ii) Prevention of harm;

iii) Fairness;

iv) Explicability.

Even though these principles suggest a guidance for solutions without a doubt, they stay as notional ethical formulas. In order to achieve Trustworthy AI, there should be particular requirements along with principles (The High-Level Expert Group on AI (AI HLEG), 2019a). According to the Guidelines, the non-exhaustive requirements are listed as follows:

i) Human agency and oversight, consisting of fundamental rights;

ii) Technical robustness and safety, consisting of resilience to attack and security, fall back plan and general safety, accuracy, reliability and reproducibility;

iii) Privacy and data governance, consisting of respect for privacy, quality and integrity of data, and access to data;

iv) Transparency, consisting of traceability, explainability and communication;

v) Diversity, non-discrimination and fairness, consisting of the avoidance of unfair bias, accessibility and universal design, and stakeholder participation;

vi) Societal and environmental wellbeing, consisting of sustainability and environmental friendliness, social impact, society and democracy;

vii) Accountability, consisting of auditability, minimisation and reporting of negative impact, trade-offs and redress.

Although all requirements are equally essential, during application, tensions among the four principles may appear. Therefore, the context and possible tensions will need to be considered crosswise different domains and industries (The High-Level Expert Group on AI (AI HLEG), 2019a). Furthermore, after the principles and requirements for Trustworthy AI are determined, technical and non-technical methods need to be completed in order to achieve the requirements. To determine the method, in these Guidelines, the entire life-cycle of Trustworthy AI given in Figure 2.3 has been suggested.

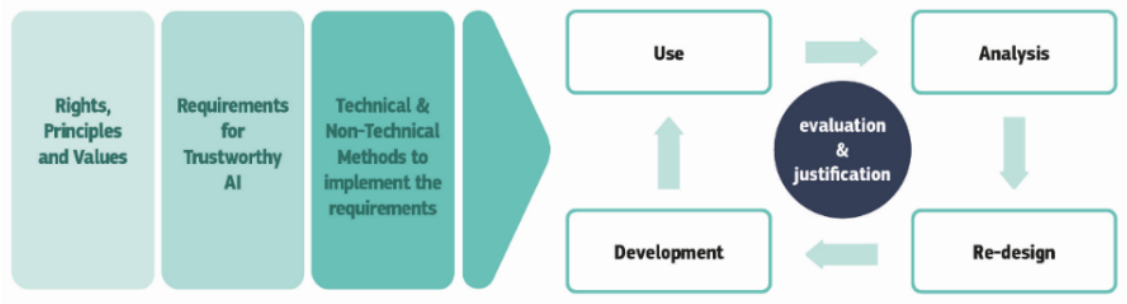

Figure 2.3: Framework proposed by the EC.

Technical methods are in a mature phase depending on the method including the design, development and use phases of an AI system. Technical methods are architectures for Trustworthy AI, ethics and rule of law by design, explanation methods, testing and validating, and quality of service indicators. However, non-technical methods are still in the immature phase consisting of regulation, codes of conduct, standardisation, certification, accountability via governance frameworks, education and awareness to foster an ethical mindset, stakeholder participation and social dialogue, diversity and inclusive design teams (The High-Level Expert Group on AI (AI HLEG), 2019a). To implement Trustworthy AI, the Guidelines indicated the importance of using both technical and non-technical methods together. Moreover, involving stakeholders during the AI system's life cycle is also considerable in order to increase their awareness towards Trustworthy AI (The High-Level Expert Group on AI (AI HLEG), 2019a). Considering the principles, requirements and methods, the EC has highlighted the adoption of a Trustworthy AI by giving a non-exhaustive assessment list for the stakeholders. Hence, the stakeholders can use this list during the development, deployment, and the use of their AI systems, however, the list should be adapted to the specific use case (The High-Level Expert Group on AI (AI HLEG), 2019a). As a result, in these Guidelines, the life-cycle of Trustworthy AI framework and the assessment list are given as a pilot, and all stakeholders are invited to identify continuous requirements, evaluate solutions, and ensure improved outcomes.

Recommendations for Trustworthy AI

In Ethics Guidelines for Trustworthy AI, the EC addressed four principles and seven requirements in the concept of Trustworthy AI. After these Guidelines, in order to advance and expand them, 33 policy and investment recommendations were announced on June 26, 2019 (The High-Level Expert Group on AI (AI HLEG), 2019b). In general, maintaining a risk-based governance approach to AI, assuring an appropriate regulatory framework, embracing a holistic way of working, examining a 10-year vision with a rolling action plan, and facilitating AI ecosystems through Sectoral Multi-Stakeholder Alliances were discussed and recommended in the document.

2.5.2 A New White Paper on AI

Following the Ethics Guidelines for Trustworthy AI and the recommendations given in June 2019, the EC has recently declared a new white paper on artificial intelligence. In this paper, their desire is to document policy alternatives considering the achievement of understanding of AI and the route of risks associated with the uses of AI considering the EU values and rules (European Commission, 2020). First of all, an achieving an ecosystems of excellence and creating a unique ecosystem of trust are the main elements of the paper. To set up an ecosystem of excellence, some actions are indicated such as working with Member States, focusing on skills, SMEs and the efforts of the research and innovation community, doing partnerships with the private sector, promoting the adoption of AI by the public sector, securing access to data and computing infrastructures, and getting international aspects (European Commission, 2020). Furthermore, creating an ecosystem of trust with respect to the regulatory framework for AI had already been discussed in the Ethics Guidelines for Trustworthy AI by highlighting seven requirements. According to this paper, the key features of those requirements are indicated as training data, data and record-keeping, information to be provided, robustness and accuracy, human oversight, and specific requirements such as biometric identification. In this paper, the EC has suggested the following ideas on how the legislative framework could be advanced to reduce risks:

- Effective application and enforcement of existing EU and national legislation;
- Limitations of scope of the existing EU legislation;
- Changing functionality of AI systems;
- Uncertainty with regards to the allocation of responsibilities between different economic operators in the supply chain;
- Changes to the concept of safety.

As one of the ideas above, the limitations of scope of the existing EU legislations has been called as a key issue for the future framework for AI. The importance of the adequate and flexible description of AI has been highlighted for AI legislations in this paper. The EC has also demonstrated the risk-based approach which requires accurate criteria to distinguish AI applications from each other by putting a high-risk determination. They also have pointed out that this determination should be accurate and readable for all companies. Besides the applicable legislation for high-risk AI applications, voluntary labelling has been proposed for "no-high risk AI applications", with which parties can increase the trust of users in AI systems in terms of certain objective and standards (European Commission, 2020).

With regards the requirements, the key features, and some suggestions; to create a future clear regulatory framework, the EC is still waiting for recommendations from all concerned parties (European Commission, 2020). Additionally, Geoff Hinton, one of the AI researchers of Google, and a professor at the University of Toronto stressed an "explainability to whom?" problem of the white paper as: "Suppose you have cancer and you have to choose between a black box AI surgeon that cannot explain how it works but has a 90% cure rate and a human surgeon with an 80% cure rate. Do you want the AI surgeon to be illegal?" (Kahn, 2020). Last but not least, the EC seeks

to set up a new regulatory framework like they did in GDPR, however, they need to create such an explainable framework for all sectors.

2.6 Conclusion

This chapter presented a literature review on the applications of AI in the financial sector, its effects on the different components of the sector and several suggested solutions for these effects. An overview of the findings are given in Table 2.1.

Table 2.1: Literature review findings summary.

AI uses	**Ethical Issues**	**Suggested Solutions**
Customer support	Complexity	Adjustments on the regulations
Detecting fraudulent transactions	Customer dissatisfaction	Building readability of the algorithms
Digital banking platforms	Discrimination	Check the inaccuracies
Reduce costs	Explainability	Close the accountability gap *
Risk management	Privacy	Coordination in the company
	Stereotyped decision mechanisms	Diminish social impact *
	Transparency	Having diverse teams
		Hire new employees
		Human-in-the-loop
		Improve the systems
		Labelling impacts *
		Mitigate possible failures (FMEA) *
		New frameworks
		Orchestrating with stakeholders *
		Succeeding the transparency *

*: Suggestions from the SMACTR framework.

Specifically, clarifications are brought on how AI is being used in the financial sector during the transformational era, and its ethical effects on customers and operations. For instance, anticipating algorithms generate complexity by using the data regarding financial services and AI. After the ethical repercussions of AI, current and suggested solutions in the literature are provided regarding AI and ethics. Further, the significance of the external audits is indicated since it can be noted that independency and objectivity are not achievable for the financial companies without them. In the last section, existing regulations for AI considering ethical repercussions in the EU are presented. As a result, the proper framework and legislations needs have been indicated to solve the above ethical consequences of AI in the financial industry in the research and EC papers. Although countless suggestions and new frameworks about Trustworthy AI are discovered in the market during the thorough literature review, a specific framework for external audit companies to solve the ethical issues has not been encountered.

Chapter 3

Methodology

In this chapter, the research methodology carried out in this thesis, with the purpose of bringing answers to the research questions declared in the Introduction, is introduced. A qualitative research approach is adopted, in order to examine the research questions. This thesis mainly focuses on the current ethical issues in AI applications in the financial industry, especially banks and insurance companies. Extensive face-to-face and remote interviews are conducted to examine the peculiar insights related with these sectors.

The grounded, interpretive qualitative study has been considered in order to understand how ethical issues in AI applications are being considered by banks, insurance companies and experts in these sectors. By taking this approach, the complexity of the field was intended to be solved. The ethical issues and AI technology by itself are complex in a realistic context. Hence, instead of any other method of qualitative research, it is advantageous to make multiple interviews with the experts, who can clearly explain the challenges associated with AI and ethics. In the context of this thesis, a total of nine interviews are carried out. All nine participants are represented by letters to make them anonymous. For example, Participant A is the insurance director in one of the Big 4 accounting firms, and Participants E-I and E-II are from the same company with different functionalities. The details of the respondents and the corresponding interviews are given in Table 3.1.

Table 3.1: Interview details.

Participant	**Function**	**Company**	**Duration**
A	Insurance Director	One of the Big 4 Accounting Firms	1' 50 min
B	Compliance Advisor Ethics and Awareness	Bank & Insurance Company I	1'40 min
C	Program Director of AI	International R & D and Innovation Hub	1' 8 min
D	Head of Data Analytics and AI	Bank & Insurance Company II	1' 10 min
E-I	Head of Group Compliance & Ethics	Financial Services Company I	45 min
E-II	Head of Compliance Analytics and Tooling	Financial Services Company I	1' 25 min
F	Founder & Chief Executive Officer	Consultancy Firm Working on AI & Ethics	1' 13 min
G	Entrepreneur & Guru on Innovation	Tech Startup Company	1' 55 min
H	Chapter Lead of Data Scientists	Financial Services Company II	50 min

3.1 Interviews

3.1.1 Interview Structure

Among different views that are possible to take in the research interview method, the neo-positivist view is taken to set up a context-free truth with minimal bias (Qu & Dumay, 2011). As stated in Qu and Dumay's work, when the neo-positivist view is taken during the interviews in the direction of the objective data and information transmission, the interviewer is considered as an adequate researcher who can generate a truthful answer, and also the interviewee as a truth teller. Moreover, in the neo-positivist view, structured interviews are being used as a tool and objective explanations are assembled (Qu & Dumay, 2011). Therefore, structured interview questions are used during the interviews, together with the neo-positivist view in this thesis, in relation with the unpretentious context of questions and answers. Interview questions are prepared before the interviews. All interview questions are asked to the participants in the same order, in order to bring answers out from a list of questions. Therefore, structured interview methodology is followed in a strict manner by reading from a script to eliminate the risk of straying from the topic (Qu & Dumay, 2011).

3.1.2 Data Collection & Analysis

With the grounded theory approach, the possible solutions for ethical issues in this field have been interpreted and analysed. This theory is functional for the interpretive of meaning in the systematic analysis of data (Suddaby, 2006). The theoretical sampling and simultaneous data collection and analysis have been adopted as the key factors of the grounded theory (Glaser & Strauss, 1967). For theoretical sampling, the participants are chosen as experts from the financial industry based on their theoretical backgrounds in the field of compliance, auditing, innovation, consultancy, data analytics, insurance, and banking.

For data collection, according to Qu and Dumay, being experts in their field has been proposed for the researchers in order to collect relevant data from the interviews (Qu & Dumay, 2011). Considering this, for the interviews, several skills such as intensive listening and note taking were used along with cautious and enough preparation. All interviews have been recorded in order to fully transcribe them afterwards. This way, the interviewer can prove that he or she did not let some information out and the full transcript is also necessary to compare the answers of different respondents during the analysis phase.

In the analysis phase, as inspired by the work in (Murphy et al., 2019) which also applies the interpretive, grounded theory approach, each interview has been coded separately in terms of phrases, words, or labels that are given by the participant. The main purpose of the data analysis in this theory is constant comparison considering iterative study design (Suddaby, 2006). All interviews have been reread multiple times to anticipate similarities and differences through participants by highlighting and marking phrases and sentences. After the coding phase, the patterns have been identified considering the study of Murphy *et al.*, to make the first-order categorisation. In the first-order categorisation, further similarities and differences between categories were found. After all the codes have been discovered, this led to the creation of the second-order categories, to establish more specific theoretical patterns. In this categorisation, it is considered that the theoretical saturation is reached, i.e. no new patterns are emerging from the interviews. Another reason of using this theory besides the grounded theory approach is that in the grounded theory, the saturation is not always apparent (Suddaby, 2006). Additionally, the need for the improvement of a valid theory has been explained as: "The grounded theory techniques are inherently "messy" and require researchers to develop a tacit knowledge of or feel for when purist admonitions may not be appropriate to their research and may be ignored." (Suddaby, 2006, p. 638). Therefore, the theoretical saturation is taken in this research to make sure of the representative findings. Moreover, the indications of theoretical saturation consisting of "the repetition of information and the confirmation of existing conceptual categories" (Suddaby, 2006, p. 639), have been considered as natural and pragmatic by the theorists and practitioners. Considering these indications, the

second-order categorisation, together with the corresponding responses taken from the transcripts are given in Appendix B. These second-order categories are further grouped and analysed (Glaser & Strauss, 1967) as the third-order categories in Table 3.4. In the next section, the third-order categories are discussed thoroughly.

3.2 Findings

It has been observed that participants were curious about AI and ethics and its impacts on the business. They were willing to talk the ethical issues in AI applications in the financial industry and also to give suggestions to decrease unintended ethical consequences of AI. In line with the research questions, solutions for unintended ethical consequences of AI, the approaches towards auditing services in the field of AI and ethics, solutions to ethical issues on AI with auditing services, and suggestions for AI legislations in the EU are analysed in distinct sections below.

3.2.1 Decreasing Unintended Ethical Consequences of AI

A number of unintended ethical consequences of AI applications have already been discussed in the literature. These can also been in Table 3.2, which presents the second-order categories that are attached to the responses. During the interviews, some of the responses of the participants have recited these risks and solutions in line with the literature, yet some of them have added new suggestions to mitigate the risks. The additional suggestions are denoted with a *.

Table 3.2: Second-order categories related to the unintended ethical consequences of AI.

To decrease unintended ethical consequences of AI.
A change in the strategy or the capabilities of the institutions is needed.
A mutualisation perspective is needed.
Adaptation is needed.
AI is more powerful than Benford's Law.
AI is the solution, not the problem.*
An ethical policy is needed.
Anticipate the data.
Be aware of the risks of technology, not be too dependent on it.
Be careful on what is happening behind the scenes in your tools.
Collaboration with other departments.
Consider long term effects.
Counterbalance the current bias.
Create a library of AI solutions for readability.*
Data quality is important.
Data scientists can tell you the accuracy of algorithm, not ethics.
Define what unethical and ethical means to you.*
Fix the biases.
Hard to work with data scientists because of the different views.*
Keep a balance between explainability, transparency, and interpretability.*
Keep HITL.
Make a honest comparison between a human being and an AI system.*
Reverse the outcomes.*
Stakeholder analysis.
The risk depends on how you apply AI.
Train people.
Trust is important especially in finance and healthcare.
Understanding is important (AI, language, biases, etc.).
Usage of the customer data dilemma.

*: Additional suggestions.

Suggestions supporting literature

Some of the responses have approved the literature in the field of mitigating the risks of AI applications. Banking and insurance companies should alter their strategies and capabilities by adapting the models to environmental changes together with keeping HITL. For the AI applications, the need for an ethical policy has also been highlighted by Participant A. Moreover, the importance of understanding of AI, its language, biases etc. has also been pinpointed by Participants A, B, C, D, and E-I. In addition, Participants E-I and F have recommended that people in the organisations should be trained to control AI applications and the risks, in addition to the understanding of AI. Participants B and F have highlighted that besides understanding of AI and training people, banks and insurance companies can also make a stakeholder analysis before they use external digital platforms in order to make sure that their platforms are safe. Additionally, collaboration with other departments has also been asserted by Participants B, D, and H. Banks and insurance companies should also use their customer data properly to take away the dilemma of usage of them which has also been suggested by A, B, D, E-I, and E-II. Regarding the usage of data, Participants A, E-II and H have recommended that banks and insurance companies can take the mutualisation perspective to create a win-win situation considering the shareholders' and stakeholders' expectations, making customers realise that they gain benefits from it. Furthermore, it has been recognised that fixing and balancing the biases are also significant for the companies based on the responses. Also, the beneficiary side of AI applications when compared to Benford's Law has been approved by Participant A. As a result, based on the interview responses, considering long-term effects of the ethical consequences of AI, banks and insurance companies should consider the risks of AI and be careful on what is happening behind the scenes in their tools.

Additional suggestions to the literature

Building on what has already been proposed in the literature, some of the responses have contributed to this thesis with additional suggestions which are denoted with * in Table 3.2. Participant C has pointed a suggestion as follows: "You have to recheck that rule from time to time. Creating a library of decision things can be a good thing so it's like central list of decisions that have been programmed. Everybody can read which decisions are made.". Therefore, to increase the readability of AI in the company for the future, creating a library can be a solution. Participant B has also expressed that they are reversing the outcomes to increase the fairness by checking the outcomes and comparing them each other. For instance, they check the certain outcomes from the model (e.g. female vs. males data), then they reverse all women into men and vice versa to check the fairness of model. With this technique, banks and insurance companies can get fair results. Moreover, Participant B has indicated that working with data scientists is a hard way because of their mathematical thinking perspective. Therefore, in the collaboration with other departments, this should be taken into consideration. In addition to these suggestions, keeping a balance between explainability, transparency and interpretability is also decisive. Participant C has indicated the importance of the balancing of transparency and explainability: "In the insurance companies, you're going to use pictures from an accident to determine that somebody isn't fooled or not. These decisions are made much more accurate but they're less explainable. You have to balance all these two lines.". Participant H has supported Participant C by highlighting the importance of the interpretability together with explainability and also has indicated that not everything could be explainable depending on the context.

Banks and insurance companies also need to define what unethical and ethical means to them because ethical values are different for every company. Participant A has clearly expressed this point as an auditing company. Their clients request the auditors to detect unethical phenomena in their data. However, it is important that they should be aware of what unethical means for them:

> "Sometimes we hear from our clients "Help me detect unethical things in my data". The model itself is usually not discriminating, it's just doing what is learned out of the data. I think the big problem is that if you can define what unethical means to you.

> You need to work really hard with your organisation on defining what you believe is not ethical because it will be different for every organisation."

Together with the suggestions, some of the participants have also approached this topic by expressing that AI is the solution, not the problem. Participants A, D, E-II, and H have highlighted that banks and insurance companies are using AI to increase their productivity. They need to define their problem, rather than just blaming AI. Participant H has suggested the need for changing the mindset of blaming AI applications:

> "AI does not impact problems, it solves problems in the way you need to change your mindset and the way you interpret the results, so with AI it's possible to reduce the time to make decisions towards the customers to have a better price and also to have a better view on our obligations we have reserved."

Participants D and H have supported these points by calling a honest comparison between a human being and an automated system. Participant D has exemplified that "AI systems which are often trained based on historical decisions of human beings so in the end they will just copy the behaviour, hence make the same mistakes. We are also making decisions that are not easy to interpret applications. For example, classifying pictures are very difficult to understand. It's very difficult to understand why he decides that a picture of a cow. It is also very difficult for a human being to explain.". Participant H has approved this claim as follows: "Of course, AI makes mistakes but they will make mistakes less than the mistakes of humans.".

After the interviews, it has been recognised that in banking and insurance companies, for AI and ethical repercussions, above suggestions should be considered before judging the AI applications.

3.2.2 Auditing Services with Ethical Issues in AI Applications

In Table 3.3, the second-order categories attached to the responses of the participants are given, for the responses related with the research questions of this thesis. During the interviews, two different approaches on AI in the auditing services were identified: the controversial and the pro-active approach. These two categories are demonstrated in Table 3.4which depicts the third-order categories. Below, the third-order categories attached to the participant responses are analysed using these two approaches.

Controversial approach

Based on some responses received, the controversial approach has been discovered. The reason is that some participants have indicated their opinion on auditing services in the ethical issues of AI applications in the financial industry by approaching suspiciously. They found that this relationship is debatable because of some reasons. To give an example, based on the response of Participant B, an external ethical board can only give advices, which is not mandatory. However, it is questionable whether these advices can solve the ethical issues in AI applications in banks or not. Participant H has supported Participant B as follows: "...they could give us advice and sometimes respect over company so we could put some best practices and to put them in place. But audit itself, I don't think they would solve the problem.".

The conflict of interest is crucial for banks. They are neatly operating their outsource activities in key banking issues. Therefore, Participant B has warned about the selection of external audit companies considering the crucial issues in banking.

There can be other factors affecting the audit committee as well. For instance, softness of ethics can affect the audits and the audit committee. Participant A has explained the reason as "...it's very soft, so it's not a clear like hard requirement.".

External audits are just in the phase of giving advices on ethical issues in AI applications to the financial services companies, they are not obligatory. In the company of Participant H, they are putting in place other user-friendly tools to detect fraud with the governance rather than a specific internal or external audit team.

Table 3.3: Second-order categories on AI & ethics with auditing services and AI in EU legislations.

Auditing services with ethical issues in AI applications.	**Solutions to ethical issues on AI, together with auditing services.**	**Suggestions for AI legislations in the EU.**
Auditors can not solve the explainability, companies should do it.	Auditors should also understand the different insights.	Application of regulations is difficult to apply.
Auditors can not solve the problem.	Audits should be done fairly.	Clear guidelines, rules, and principles are needed.
Consider the conflict of interest.	Auditors should understand what AI is.	Clear guidelines and rules are needed.
Internal audits can be slow in AI and ethics audits.	External auditors should gain the view on the internal corporate policies of the clients.	Consider certain values.
Pro-active approach with external auditing services.	Having own challenger models in auditing services.	Consider the co-ordination and globalisation.
Put in place some user-friendly tool rather than internal or external audits.	Understanding each other in multidisciplinary teams.	Do not kill the innovation.
Softness of ethics can affect the audits.	Standards are needed for external audits.	EC should reflect their purpose.
	Test the tool, think the whole process.	Follow data science standards.
	Try to see hidden purposes in the algorithms.	Hard to define risks.
		Having more balanced view is important.
		Labelling can be mandatory.
		Problem is the lack of accountability.
		Regulators should guide and audit.
		Standards and boundaries are needed for the voluntary labelling.
		The difficulty is the translation of ethics.
		The white paper is so narrow.
		Understand the language, the problems, and the purposes of AI.
		Voluntary labelling can be useful for establishing the trust between parties.
		Voluntary labelling will not work in practice.

Table 3.4: Third-order categories on AI & ethics with auditing services and AI in EU legislations.

Auditing services with ethical issues in AI applications.	**Solutions to ethical issues on AI, together with auditing services.**	**Suggestions for AI legislations in the EU.**
Controversial approach	Be fair	Accountability dilemma
Pro-active approach	Appreciate	Broad-minded approach
	Having an understanding in multidisciplinary teams	Clarify the applications, simplify the language
	Own challenger models	Counterbalanced approach
	See what is hidden	Crucial need for rules
	Standards	Do not forget gloablisation
	Tailor-made audits	Give birth to innovation
	Test	Master
		Respect the values, reflect the purpose
		The reality of risks
		Voluntary labelling is debatable

Pro-active approach

After the interviews, it has been observed that external audit services can control risks in AI. Participant F has clearly indicated that internal audits can be slow in AI and ethics audits. From this point of view, Participant F has also explained the reasons of the need for external audits:

> "If you have one person looking at it, from their perspective having the audits you can get a clear view of where the big points are. If you are aware of the risk points prior to any type of situation around risks, you will be prepared so it is shifting pro-active rather than re-active in risk problems."

Participant C has agreed with Participant F:

> "I would suggest not to do it internally but asking an external person. External auditors will see the issue easier."

However, Participant C has also highlighted that the explainability of the model is the responsibility of the company, not the auditing services. Based on these responses, to get a clear view and understand the risks pro-actively, the external auditing services are needed in this field.

3.2.3 Solutions to Ethical Issues on AI together with Auditing Services

During the interviews, the solutions to mitigate the unintended consequences of AI applications have been expressed by most of the participants. While they were giving suggestions in Table 3.2, they have not specifically mentioned that auditing companies should consider them. However, when they are compared with the responses given in Table A.4, which contains the responses on the solutions to ethical issues on AI together with auditing services, some similar responses have been observed. Examples to such responses are the importance of understanding of AI, the fairness aspects, the need for standards and the changing techniques for AI applications.

Be A Hosstt Framework for Auditing Companies

After the second-order categorisation of the responses in Table 3.3, they are grouped together to form the third-order categories as indicated in Table 3.4. These third-order categories are fitted

into a framework called "Be A Hosstt", as illustrated in Figure 3.1. This framework would be useful for auditing companies in their AI and ethics audits. This framework consists of eight essential elements –which complement each other– that external auditors need to keep in their mind. The name "host" is related to the fact that they should act as hosts and moderators. For instance, they can detect ethical issues of AI applications in financial services by acting as controllers.

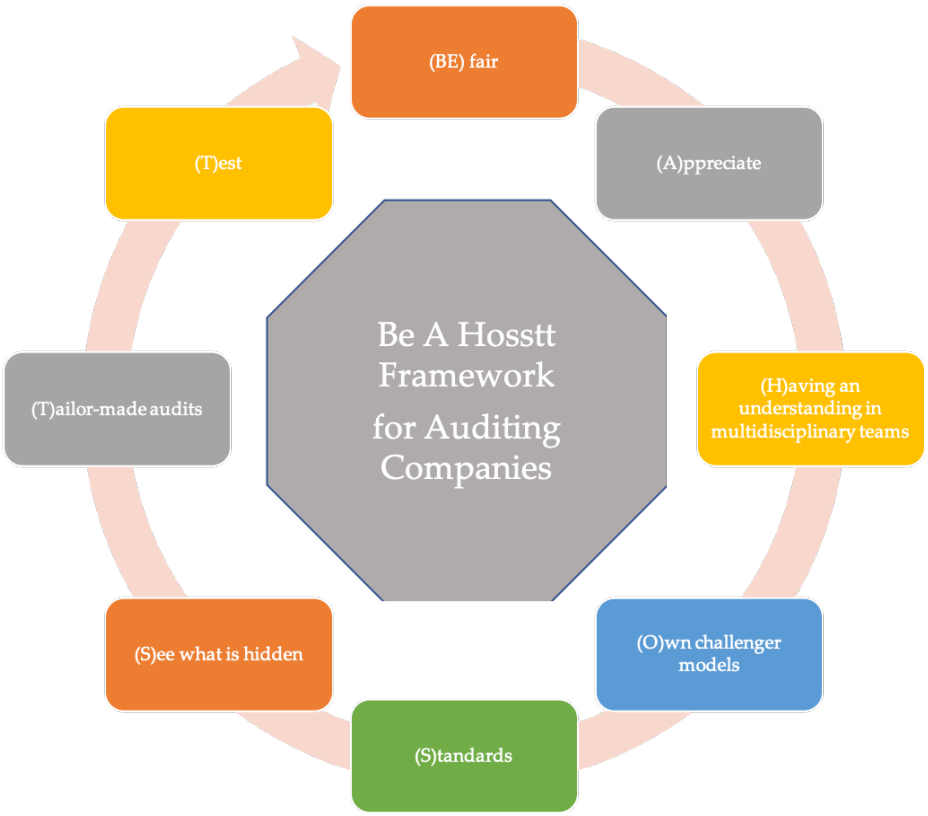

Figure 3.1: Be A Hosstt Framework for auditing companies.

Be fair.

The external auditors should always keep fairness in their mind. Their main purpose is to be independent and objective in order to give their opinion in a fair way. Participant H has highlighted this point as: "All audit is good as long as it is done in the fair way, in the collaborative way.". Participant E-II has also indicated that they should be independent.

*A*ppreciate.

Appreciation is a key element in AI applications. Hence, auditors should understand what AI is and also its different insights. Participant E-II has recommended this understanding with different insights in the following way:

> "...They should be aware that they should be able to understand what AI is about data, what is the sensitivity of data, and how may be putting together data which in isolation or not sensitive."

Participants C, G and H have also approved the importance of understanding every single part and the purpose of AI. Participant C has pointed out that this appreciation should be done by having different insights:

> "...You have to make sure that you have access internally to the right people who can give you the answers because you will need to have the different insights."

Based on these responses, auditors should appreciate and internalise what AI means and what the purposes of models and audits are, together with different insights from the right people.

_H_aving an understanding in multidisciplinary teams.

As important as understanding of AI, understanding people is also needed in external audits. The reason is the different meanings of trust for everybody. Participant C has given this suggestion as follows: "We all use the word "trust" but we all mean something differently. That's why, also in audits, when even if you make an effort to do for multidisciplinary team, you have to make sure that people understand each other." Therefore, having multidisciplinary teams are needed in external audits together with understanding each other. With these teams, auditors can combine several skills and understandings during the audit process.

_O_wn challenger models.

Having own challenger models in auditing services is crucial to get the results that auditors expect. Participant A has highlighted that they have their own challenger models such as right pricing models to check the pricing at the insurance companies in order to make sure that their results are in line with what they expect.

This element is complementing the "Appreciate" and "Be Fair" elements of the framework. After external auditors evaluate the results with their own models, they can understand companies' models to clarify the errors further. If there is an error in the companies' models, auditing companies can detect them with their own challenger models. As a result, objectivity and independency will be reached by checking the models with the different model as well.

_S_ee what is hidden.

Participant G has suggested that trying to see hidden purposes in the algorithms should be the main purpose for the auditing companies. After they try to figure out unseen parts of the algorithms with their multidisciplinary teams, they will gain trust from their clients.

_S_tandards.

Standards are needed for external audits to audit their clients based on the rules. Participant A has highlighted this need as: "You need to define what exactly your rules are. As an audit firm, you need to go to rules and standards." This is supported by Participant E-I as follows:

> "For example, there is an association of internal auditors, it is producing standards to support auditors in doing their job. Having similar standards to audit AI is certainly something that would be useful for internal audit or another lines of defense."

Participant C has approved the other participant by also recommending a guarantee label for AI applications which can be given by the auditing companies as in other sectors such as food and medical. For the auditing companies, they need to have standards which are set by the authorities to give a guarantee label to the companies. To be complied with the standards, external auditors can apply these standards on their own challenger models to reach a good quality of work to mitigate the unintended consequences.

_T_ailor-made audits.

Every company has different corporate policies and values, hence, external auditors should gain a view on the internal corporate policies and values of their clients. Gaining a view for the company values has been expressed by Participant B in the following way: "I think auditing should be linked

to the company values. It has to be tailor made." Participant E-II has approved this by giving an advice on considering companies' internal corporate policies in the audits:

> "In my opinion, the audit should also have a view on what are the internal corporate policies of their clients because there's a second side of the coin."

Considering these responses, external auditors should adjust their methods in their challenger models based on corporate policies and values of the companies to make tailor-made audits together with the consistent standards.

***T*est.**

A testing tool by thinking the whole process of AI should be one of the basic elements of all auditors. They need to check whether their tools are working or not. Participant A, as an external auditor, has recommended this suggestion to internal auditors:

> "If you are an internal auditor, the big question you always need to answer, is my tool working? You need to even test your tool that it's really working to do an audit because the worst thing you can do is to use a tool that is not catching any risks that you want to detect. Secondly, if you're using that tool like a design tree to find something, in audit, one of biggest thing is always you need to come up with an audit approach. So you always need to go through the whole process of thinking "What is the approach for me about the risks I want to cover?"."

Participant D has approved this response as follows: "I think it's more important to understand and to be very well focused on testing the capabilities of a model first.". Considering these suggestions, all auditors (in this framework, external auditors) should always check their "Own challenger models" and tools, and consider the whole process by asking questions themselves about their aim to figure out the risks properly.

3.2.4 Suggestions for AI Legislations in the EU

In this section, the third-order categories associated with the third research question are given, together with the corresponding quotes from the participants.

Accountability dilemma.

The participants have pinpointed that there is a problem with the lack of accountability. Companies do not know where to put the responsibility and accountability. Also, they do not have any idea about how to cope with the accountability problem if they face with it. Participant F has highlighted the accountability issue:

> "The problem is due to the lack of accountability. You have to decide who is going to take the responsibility for it?...The problem is no one is willing to take responsibility."

Participant C has supported that the accountability of AI application does not belong to data scientists since they are not aware of the possible ethical issues produced by their applications. Participants A and B have approved them by asking the same question as "Who is accountable for AI applications if there is an issue in it?". In light of the given responses, EC should solve the accountability issue of AI applications with the legislation.

Broad-minded approach.

The new white paper is not enough from all perspectives. The missing parts and the narrowness of the legislations should be realised by EC based on the responses of Participants B and C. The missing part has been figured out by Participant C:

> "It's just focusing on one end application for example a risk application or fraud application. The world is where you're going to have two autonomous vehicles meeting each other. For example, if they have an accident, you basically have two high risk applications competing with each other. The white paper doesn't say anything about the interaction between two high risk applications so it tends to take a human versus a system whereas we're going more and more into a world where we have a system against the system and it's nothing mentioned there about the trustworthy part."

Participant B has invited EC to broaden the existing legislation as well. Accordingly, EC should broaden the existing regulations and realise the missing parts in the paper. They should consider the possible danger of interaction between two high risk applications as well since it has not been mentioned in the new white paper yet.

Clarify the applications, simplify the language.

Based on the responses of Participants B, C, D, E-I, and F, it has been observed that the application of regulations is complicated and not straightforward even for large and small companies under the complexity of the world, technical and financial constraints. Participant F has indicated that these principles are a good initiative, however, the EC should simplify the language of application:

> "From an ethical aspect, it is great to list these principles of EU framework. It was good for step because you need a kind of code. What would you do after that? How do they be applied? These are the issues."

Additionally, as the difficulty of the softness of ethics for auditors has been analysed earlier, its translation is also not simple. Therefore, as Participants A, B, and D have pointed out, the implementation of the AI and ethics principles can be struggling for the companies, hence the EC should find a way to adapt these principles by making them understandable. Therefore, the EC should solve these complications firstest with the mostest to help companies.

Counterbalanced approach.

Having a more balanced view is important for EU legislations in AI and ethics. Based on the response of Participant D, EU legislations should not be overcritical to AI applications considering both IT-based and human-based decision systems. A balance is needed in this field because these two systems need the same level of control and objectivity. Participant E-II has supported this approach as:

> "At least you have some transparency and the collaborative work around the topic and having the view of multiple people allows you to have a more balanced view and, hopefully have more balanced constraints in the end. Therefore, EC should approach to AI applications and its legislations with the balanced view."

Crucial need for rules.

During the interviews, the need for rules, standards and principles has been indicated by most of the participants. In the shortest time, EC should announce the legislations of AI which includes rules, standards and principles in a concrete and clear way. Participant D has highlighted this crucial need. Participant A has supported Participant D by highlighting that auditors can move more strategically in their AI and ethics audits considering the EU standards:

> "There are lots of frameworks in the market. This will keep on evolving, and it's important to follow that up. I think everything is on the table now but it's really waiting the European Union as a power to really say "These are the principles..." and if there is a standard clear rules, I think auditors can do something with that and they can check the rules against the application. So that is clearly on the strategic piece.

> There's one central database of the government that will help the whole sector and that's clearly a point to improve the control function."

Participant E-I has also approved the importance of the need for the standards to help business and auditors:

> "In order to add value to the business, benchmarks are made available to help business to help auditors etc. but it needs to be more concrete. Also they should not leave these principles as too much room for interpretation."

Additionally, Participant H has recommended the following data science standards in order to protect tools:

> "I think the key will be to follow the standard of data science. If it is not a data scientist playing with the tool, it means that someone who doesn't have the knowledge and hasn't been trained correctly will use the tool."

Considering above responses, the EC should take actions to create proper standards and rules for the business and auditors. They can also follow data science standards while creating standards.

Do not forget globalisation.

The EC should regard the co-ordination and globalisation in the legislations of AI and ethics. Participant B has marked the importance of globalisation:

> "You have the globalisation so you really have to take into account it...you also see now for example in GDPR, it is becoming more and more important so in other parts of the world. You see also in the US."

Participant E-I has indicated that EC should collaborate and communicate with other countries before they release the rules on AI and ethics:

> "It will be very important to liaise with other like the US, India or China in order to get as much alignment as possible."

Give birth to innovation.

While putting the legislations, sometimes innovation can be damaged because of the limitations. In the legislations, the flexibility should be considered. Participant E-I has drawn attention to the importance of the appropriateness of the innovations together with the legislations:

> "We have to put in place regulatory framework that would avoid the abuse of consumers data for AI but at the same time we have to be careful not to develop legislations that would be considered as being counterproductive versus the appropriate of innovation."

Participant G and H have approved that the innovations should not be killed by lots of limitations. Instead of killing the innovation, the legislations should be done in a smart and interesting way. Hence, the EC should put standards, principles and rules towards AI and ethics, however, while they are creating them, they should be taking the flexibility of the innovation into consideration as well.

Master.

The EC should be a master to control the unintended consequences of AI applications. First of all, they can start being a master by understanding the language, problems, and the purposes of AI. Participant G has called attention to this understanding:

> "I'm sure through privacy some of those questions are already coming but maybe it doesn't cover enough of the purpose. For the regulators, I would focus on the purpose of why we collect data and why we build models."

The importance of defining and understanding of AI has been acclaimed by Participants A, D, E-II and F. Participant F has also indicated that regulators should educate themselves first to understand the problems of AI. Afterwards, the EC can audit and guide companies to make sure the correctness of their application based on Participant G's additional response on this topic. Therefore, the EC should first educate themselves, then understand what AI is, and audit and guide companies as a final step.

Respect the values, reflect the purpose.

In the EU, there are certain values which are different from other countries. Therefore, Participants A and B have pinpointed that the EC should consider these values by taking a global perspective. These participants have exemplified that the EU sometimes takes into consideration these values by not not allowing face detection in the streets which is seen in China. They should always set the policy in order to be submissive with the EU values and culture. Based on the responses of Participants B and G, whilst the EC is respecting these values, they should also reflect their purposes by demonstrating that the EC is trying to help companies with the legislations.

The reality of risks.

Participant C has demonstrated that in the models, there can be errors which can impact the definition of the risks considering the classification of no-high risk applications which is indicated in the new white paper of EC. Participant C has also exemplified the possible complexity as follows:

> "They look at single applications. It's like in a plane, every individual component might be perfect but in airline crashes and what you see, it's because the threshold of one application was good enough to pass information to the next application where it was good enough to pass the next application and the final application fails but it's because of the very early one. I find it risky to classify it like this because the world is a little bit more complex."

Participants D, E-II, F, and G have confirmed that the description of the risk criteria should be described by the EC considering the transparency, interpretability, and context. Participant F has also expressed that there should be more stricter rules in this description:

> "My first question to them would be how do you define high risk, one of the things with AI is you can create an AI application and use it in a low risk situation but as soon as it's taken out of that context it would be considered high risk so how do you define that in the first place? I think, yes, it makes sense to have a bit stricter rules in terms of high risk."

Consequently, the EC should define the high risk and no-risk applications based on some criteria and make them more rigorous.

Voluntary labelling is debatable.

In the last new white paper, a voluntary labelling for no-high risk AI applications proposal has been released. Some of the participants have responded that as long as it is mandatory, it can work. Participant E-II has expressed that companies can be forced to think in the very early stages of the application if this labelling is mandatory. Hence, it can be beneficial for companies to consider it in the initial phases. Participants F and H have indicated that there should be standards and boundaries for the voluntary labelling in order to protect fairness and random labelling.

On the one hand, Participants B, C, and E-I have pointed out that voluntary labelling will not work in practice considering the cost, existing many norms in the organisation, and the complexity. On the other hand, Participant A has highlighted that voluntary labelling can be useful for establishing the trust between parties:

> "In Belgium, you can outsource the activities of calculating the payroll. The problem is a company that you don't see what's happening in the centre. So what the solution there is that they are audited every year or multiple times a year based on a list of criteria and then they issue reports to your company so you know that they are taking care of my payroll but I have a report that everything is really functioning like it should be on there. And so that gives them trust between these two parties. That could help you in establishing relationships between parties in the market. I think voluntary labelling would help and I think we've already seen this type of certifications in other spaces like payroll services or more."

As a result, voluntary labelling has been discussed among participants. The EC should consider the standards, boundaries and its obligation to get results from the labelling. Then it can enhance the trust between parties as a kind of certification which can help auditors as well.

3.3 Conclusion

In this chapter, the methodology of this thesis –qualitative research– has been indicated. Afterwards, the interview details have been shown in terms of the functionality of participants, their sectors, and the durations of the interviews. The transcripted interviews have been coded in three different levels, and the second- and third-order codes are introduced in this chapter. Following the completion of coding, an analysis is conducted. As a result of the analysis, a new framework with the name "Be A Hosstt" has been proposed in relation with the second research question. The results obtained from the analyses are further used in the discussions in the next chapter.

Chapter 4

Discussion

The research questions of this thesis examine the relationship between ethical issues in AI applications and auditing services in financial services, solutions for external audit companies to decrease the ethical consequences of AI and suggestions to EU legislations in the direction of this field. During the interviews, it has been observed that financial services companies use AI to gain a competitive advantage. However, they are also aware of the potential risks associated with it, as warned in the literature. They have already taken these risks into consideration and have made suggestions to decrease these adverse effects of AI applications.

As indicated in Table 3.2, some suggestions from the literature have been recited by the participants. For instance, in order to gain their customers' trust, Participants A, C, and E-I have recited the importance of the change in the strategies and capabilities towards AI systems based on their main customer segments considering the digital transformation era in line with the suggestion of "improve the systems" from the literature. Moreover, the need for adjustments on the regulations suggestion has also been supported during the interviews. In the literature, chatbots and NLP have been considered as risky applications which are creating biases and decreasing the trust between companies and customers. However, by keeping HITL, these consequences can be prevented based on the responses from the interviews, supporting the ideas of Green & Chen (2019).

Additional Suggestions to the Literature

In addition to suggestions that are already present in the literature, the participants have also expressed the following additional ones which can also be found in Table 3.2:

- Hard to work with data scientists because of the different views during the collaboration with other departments,
- Creating a library of AI solutions for readability,
- Defining what unethical means to companies,
- Keeping a balance between explainability, transparency, and interpretability,
- Making a honest comparison between human-being and AI system,
- Reversing the outcomes.

In the literature, there are some missing parts considering these suggestions. To illustrate, in the report of MIT SMR Connections (2020), the importance of explainability has been highlighted to give the reasons to the stakeholders about approved or unapproved insurance applications. Additionally, in the pricing policies, transparency is a key element for the insurance companies to explain their decisions rather than blaming the computer (Hall, 2017). Participants C and H have pinpointed that there should be a balance between transparency, interpretability and

explainability. Explainability by itself will not make sense since sometimes some decisions are less explainable but more transparent. Therefore, balancing is significant regarding the context of the application.

As a result, companies should anticipate these mistakes before they occur rather than just blaming AI mechanisms. AI is the solution, not the problem for these risks, as long as companies understand what AI means and what they want to solve.

The Relationship of AI and Ethics with Auditing Companies

In the literature, significant suggestions and frameworks have been proposed for AI applications. However, no specific suggestions or frameworks for external auditing companies were found that are specifically designed for AI applications. More specifically, the existing frameworks are mainly intended for internal audits. The importance of the relationship between internal audits and AI and ethics has been figured out in the literature as it can be seen in SMACTR framework which combines the structure of internal audit and engineering processes (Raji et al., 2020). Based on the responses received, the reason behind this importance might be that there are no specific standards or rules for external audit companies. Considering other factors such as softness of ethics, conflict of interest for banks and the existence of user-friendly AI tools, some participants have approached external auditing companies with a doubt. This approach is determined as the "Controversial Approach" in this thesis. Per contra, some participants have expressed that the external audits would view more clearly and understand the problems intensively. This approach is considered as the "Pro-active Approach" since the external audits anticipate and control the unintended risks of AI beforehand and objectively. Internal audits can be slow and re-active in the detection of these risks due to the fact that external audits have an analytical power to detect risks (Muñoz-Izquierdo et al., 2019).

According to Raji et al. (2020), internal auditors have the capability to access internal information easier than external audits. Hence, external auditors can use this information in the detection of biases and risks. Similarly, for financial information audits, the CA system has been proposed by Chan & Vasarhelyi (2011). The CA system has highlighted the importance of both internal audits and external audits. In this system, in the implementation phase is being completed by internal audits, afterwards external audits give the certification to their implementations by acting independently. The logic of this system could be applied for AI and ethics audits. Accordingly, internal audits can implement the organisational AI principles in engineering processes and implement reporting requirements properly, and use the audit results for the necessary organisational changes as a suggestion (Raji et al., 2020). External audits can detect further biases and risks and perform as independent certification providers of these audits. In line with this suggestion, if the EC puts the proper standards for auditing companies, auditing companies can give certificates to the companies independently in specific periods. By doing so, external audits would no longer be in the advisory stage. As a result, they would detect the biases and risks of AI pro-actively, and thus, the aforementioned unintended ethical consequences of AI would be decreased by this control mechanism.

Solutions for Ethical Issues of AI Applications in Financial Services with Auditing Companies

Banks and regulators should increase their capabilities and understanding of their practices with AI to improve their financial performance by establishing a risk-aware culture in the organisation (Fourie & Bennett, 2019). Additionally, the SMACTR framework has been proposed by Raji et al. (2020) for internal audits to build up AI systems.

Similar frameworks, such as the SMACTR framework, have not been found in the literature for external audits. Therefore, this thesis has contributed to the literature by suggesting a framework for auditing companies which is given in Figure 3.1. The "Be A Hosstt" framework is a principle-based framework, not a technical framework as SMACTR for internal audits. There are eight elements which are compatible with each other in this framework. First and foremost, based on the

responses of Participants H and E-II, external auditors should not forget to be fair and independent as a key element. Second principle which is similar to the logic of SMACTR framework's scoping phase, external auditors should always understand the purpose of the models, audits, and AI. After the appreciation, external auditors should also "understand" each other in multidisciplinary teams to increase the quality of work, as an additional suggestion to the literature. Moreover, there are several frameworks and models in the market, however, auditing companies should have their own challenger models as in the company of Participant A. After having own challenger models, auditors should try to see what is hidden in the models as a main purpose with their multidisciplinary teams to protect themselves from blindness towards the detection of the risks. To do such external audits, standards and rules are needed. Thus, auditing companies would consider the standards and give certifications to banks and insurance companies as an independent certifier. However, in their audits, they also need to consider company values and policies by tailoring their audits together with the specific standards. After they take these principles into consideration and complete them, as a final principle, they should test their tools by considering the whole process. As a result, this principle-based framework would be beneficial for auditing companies in the detection of ethical issues of AI applications in financial services by acting as a controller and an independent certifier.

Comparison of the "Be A Hosstt" Framework with the EU Guideline

In order to increase the credibility of the proposed framework, a comparison is carried out with the EU Guideline as a framework for Trustworthy AI. The results of this comparison are given below separately for each element of the framework.

- It has been observed that fairness, one of the ethical principles in the EU guideline, is crucial for AI applications in order to ensure that individuals and groups are not affected from unfair bias and any discrimination. This corresponds to the "Be fair" element of the proposed framework, the difference is that fairness is considered for external auditors (not for AI models) in their AI-related audits. As human beings, external auditors should consider objectivity and independency in this element.
- In the EU guideline, as a technical method, explanation methods have been indicated as Explainable AI (XAI) to understand the system and to find solutions. This is also considered in the "Appreciate" element of the "Be A Hosstt" framework, that auditors should understand what AI is with the different insights.
- As it has been proposed in the EU guideline, diversity and inclusive design teams take an important part of the non-technical methods in terms of different perspectives, needs and objectives. This is also considered in the proposed "Be A Hosstt" framework with the additional suggestion "Having an understanding in multidisciplinary teams". Diverse and inclusive design teams are crucial, however, people also need to understand each other within the team to combine their skills.
- The "Own challenger models" element of this framework has not been specifically proposed in the EU Guideline for Trustworthy AI framework as a requirement or any other line. Especially, in auditing services, the auditors need to have their own models to ensure the best results from them, thereby increasing the credibility of their services.
- From the requirements list of the guideline, the avoidance of unfair bias has been highlighted under the "diversity, non-discrimination and fairness" requirement. Considering this requirement, this coincides with the "See what is hidden" element of the proposed framework. External auditors need to see hidden purposes in the algorithms to avoid unfair bias.
- Besides, as one of the non-technical methods, standardisation (like ISO standards) has been indicated as an ongoing basis in the EU guideline to encourage ethical conduct in the decisions. This is in line with the suggestions proposed in the "Standards" element of this framework, since external auditors need to have standards to give certification.

- In the EU Guideline, as another non-technical method, giving certifications including standards which are compatible with different contexts in terms of the industrial and societal standards has been indicated. This method is partially considered in the "Be A Hosstt" framework. In order to be in line with the "Tailor-made audits" element, the external auditors can design tailor-made audits regarding the company values and policies, as every company has its own values and policies.
- As one of the technical methods from the EU Guideline, testing and validation of the system should be considered by companies as early as possible to safeguard the consistency of the outputs with the results. This coincides with the "Test" element of the proposed framework. External auditors should make sure that their tools are performing well.

This comparison shows that the proposed "Be A Hosstt" framework shares many characteristics with the EU Guideline as a framework for Trustworthy AI. Except the "Own challenger models" item, all the items within the proposed framework are directly or indirectly related to some elements of the EU Guideline. However, this element has not been considered as requirement or principle in the EU Guideline. Last but not least, the comparison above and the "Be A Hosstt" framework can also be taken as a sample by the experts to understand the application of EU regulations. For instance, it can be considered for one of the suggestions for AI Legislations in the EU "Clarify the applications, simplify the language".

The Improvements of AI Legislations in the EU

In Trustworthy AI, High-Level Expert Group in the EC has indicated several principles such as human agency, technical robustness, privacy and data governance, transparency, fairness, social impacts, accountability etc... EU has also proposed the framework given in Figure 2.3 which highlights that the technical methods should be combined with non-technical methods. In line with the EU Guideline, the need for the principles, the lack of accountability, and the difficulty of translation of ethics have been observed during the analysis. However, additional suggestions have also been discovered. For instance, EU should figure out the accountability issue by broadening their approach since there are some missing parts in the legislations. For instance, the interaction between two high risk applications that compete with each other has not been indicated in the new white paper, based on Participant C's response. Additionally, as Participant D has indicated, the EU should not be too critical. They should balance the AI applications and the legislations since both IT-based and human-based decision systems need to be controlled considering the same level of objectivity.

Considering "A New White Paper on AI" which is backing up Trustworthy AI, besides respecting the European values, they should properly reflect on their purpose by showing that they are there to help companies rather than limit them. They should be masters in this field by educating themselves, auditing and guiding companies. While they are doing it, they should give birth to innovation since sometimes legislations can limit the innovation and kill it. Furthermore, the definition of risk is not clear and accurate in the paper, therefore, they should create standards and boundaries for them based on the criteria. Additionally, based on the responses received, if the labelling system is mandatory, it can increase the performance of the system and protect unintended ethical consequences of AI in the early-stages of applications. If it is not obliged to do so, doing voluntary labelling can also improve the trust between company and regulators. The EC should examine these factors in the voluntary labelling, and also above factors in the legislations for AI. Last but not least, they should take an action as soon as possible since banks, insurance companies and others as well seek to get clear regulations on this field.

In conclusion, most of the participants have declaimed the lack of accountability, crucial need for rules considering globalisation and innovation by respecting the European values, solving the complication of AI applications, criteria for high-risk AI applications, and the difficulty of translation of ethics in the EU legislations. Along with them, the EC should not forget to broaden their approach, accurately reflect their purposes, put some flexibility in the legislations, and reconsider the voluntary labelling proposal again.

Chapter 5

Conclusion

In this chapter, the conclusions of this thesis are given. The conclusions of this thesis can assist external auditors, banks, insurance companies, or financial services companies in AI applications considering the ethical risks. Following the conclusions, recommendations are given for the future work.

5.1 Conclusions

This thesis aimed at giving new insights about designing the solutions for ethical repercussions of AI applications together with auditing services. Additionally, this thesis also focused on AI legislations in the EU with the purpose of giving suggestions.

During the literature review, it has been discovered that there are numerous unintended ethical consequences of AI and solutions for them, such as AI frameworks and toolkits for internal audits and companies in general. However, no specific relationships of AI and ethics with auditing companies or associated frameworks have been discovered. Following the literature review, a qualitative study –by adopting the grounded theory approach with theoretical saturation– is carried out by conducting interviews with experts in the field of AI and ethics.

Regarding the first research question "What is the relationship of ethical issues with AI applications in financial services (e.g. banks and insurance companies) with auditing services?", two approaches have been observed in the responses of the participants: controversial and pro-active. In the controversial approach, participants mentioned that they approach audit services with a question in their mind, since the auditing services for this topic are not mandatory. Additionally, other factors affect their responses as well, such as softness of ethics, conflict of interest for banks, and better AI tools in the detection of the risks of AI. However, it should not be forgotten that AI tools are designed by a human being as well. Therefore, to strengthen the control mechanisms, auditing services should be necessary in this field, defining the pro-active approach. External auditors can act as certifiers by approaching AI applications objectively and independently. They also can work collaboratively with internal audits, in order to increase the speed of the process. Based on these two approaches, the relationship between auditing services and AI and ethics in financial services would be improved, if auditing services are made mandatory by taking away the doubt in companies' mind for auditing services' advisory stage. Based on the responses received, the participants have recommended additional suggestions to mitigate unintended ethical consequences of AI. Even though they have not been specified for auditing companies, it is crucial that auditing companies take these suggestions into consideration since some of them have also been suggested for auditing companies.

To remark the second research question "How companies in the financial industry can solve their ethical issues, related to artificial intelligence, together with auditing services?", a new framework "Be A Hosstt" for auditing companies has been proposed. Auditing companies could consider the elements in this framework in their audits on AI applications. It would be beneficial for them to act

as moderators and certifiers. Last but not least, the "Be A Hosstt" framework has been compared with the EU Guideline for Trustworthy AI framework. In the comparison, it is observed that almost all elements in the proposed framework correspond to some principles and methods in the guideline. The only exception is the "Own challenger models" element. In the EU Guideline, this element has not been specifically highlighted as a principle, method, or a requirement. Therefore no relationship has been found for this element. Overall, the EU Guideline has supported auditability as follows "Evaluation by internal and external auditors, and the availability of such evaluation reports, can contribute to the trustworthiness of the technology." which can be useful for auditing services in their works.

For the last research question "What could be improved in AI legislations in the EU?", it has been observed that most of the participants have recited the problems on the legislation that can be found in the papers "Ethics Guidelines for Trustworthy AI" and "A White Paper on Artificial Intelligence". Apart from the same problems, the participants have invited the EC to broaden its approach by considering other factors such as the interaction of two high-risk applications with each other. The EC should also indicate its aims properly, consider the flexibility, and overview the voluntary labelling for no-high risk AI applications one more time. The reason behind this is that these factors have not been indicated in the legislations clearly and this causes some questions in the companies' approaches. To better understand how to apply the EU regulations, the "Be A Hosstt" framework can be considered as a clarification by the experts regarding its comparison with the EU Guideline. Lastly, for the "Be A Hosstt" framework or any other framework to be efficiently used by auditing companies, the EC should put standards on AI legislations and consider the certifications which can be given by auditing companies.

5.2 Implications for Future Research

The following points of future research have been indicated, which can take the findings of this thesis even further. External audits can bring a solution to unintended ethical repercussions with their independent and objective works. However, due to the complexity of the technology, they can not solve these problems themselves. The technology always needs adaptability, companies should adapt themselves based on what they want to solve, hence, the future research can go deeper to propose new suggestions. Secondly, the literature about the relationship between ethical consequences of AI applications and auditing services is limited. For this reason, it offers further research opportunities. Additionally, the relationship between auditing services and AI applications can be more extensively revealed by conducting more interviews. Lastly, effectiveness and validity of the proposed framework in this thesis can be tested, and a technical perspective can be brought to this framework.

Appendices

Appendix A

Interview Questions

In this Appendix, the questions used in the interviews that are carried out in the scope of this thesis are presented. The purpose of the interviews were to understand the auditing system of artificial intelligence in banking and insurance companies, and how these companies are coping with possible issues by themselves as well as in collaboration with auditing companies, regarding the biggest ethical issues in financial services. The interviews were conducted with experts from different companies in this field, particularly financial firms (e.g. banks, insurance companies, and auditing companies) and non-financial firms (e.g. consultancy firms working on AI & ethics).

The interview questions are grouped with respect to the research questions that are answered in the thesis, and are given in Tables A.3, A.4 and A.5. Introductory questions that are used for ice-breaking purposes are given in Table A.2. Optional questions are denoted with the superscript "O" in the tables given in this Appendix. As some questions are specifically aimed for certain types of companies, the type of company that each specific interview question is applicable for is also given in the tables. For convenience to the reader, the research questions are also given in Table A.1.

Table A.1: Research questions covered in this thesis.

Number	**Research Question**
I	What is the relationship of ethical issues in financial services in AI applications (e.g. banks and insurance companies) with auditing services?
II	How companies in the financial industry can solve their ethical issues, related to artificial intelligence, together with auditing services?
III	What could be improved in AI legislations in EU?

Table A.2: Introductory questions.

Number	**Questions**
I	To what extent are you applying AI to your business?
II	In the literature, for example in the IEEEXplore database, there are 200,468 publications regarding "Artificial Intelligence". The number of publications regarding "AI and Ethics" is only 302. What do you think about AI and its relation to ethics in the financial industry? Source: (https://ieeexplore.ieee.org/Xplore/home.jsp)

Table A.3: Questions related to Research Question I.

Part I: Firm's Experience		
Number	**Questions**	**Applicable For**
I	Which types of methods are used in the monitoring of the AI applications in banks that create corresponding ethical issues?	Both
II	What is the most important competitive advantage that your company/financial firms gain from AI technologies?	Both
III	What types of ethical issues have you faced? How did you approach them?	Both
Part II: Literature		
Number	**Questions**	**Applicable For**
I	What could be the effects of decision mechanisms regarding ethical perspective?	Both
II	What might banks do to secure the information collected from AI applications (e.g. from the chatbots)? How do you approach Natural Language Processing (NLP) in the financial services?	Both
III	For customer support and front-office purposes, some banks are using Kasisto's KAI -a digital assistant-to help their customers by way of making international transfers and blocking credit card charges. For fraud protection and middle-office purposes, banks are benefiting from the digital platforms in order to manage risk and safeguard from fraud, anti-money laundering initiatives, and know-the-customer identity confirmation. Do you see any potential threats from the digital platforms regarding AI and ethics?	Both
IV [O]	PayPal advanced their cybersecurity services by keeping up their fraud rate considerably low, at 0.32 percent of revenue. How do you think financial institutions can advance their AI applications considering ethical issues, to decrease their fraud rate and increase their knowledge?	Auditing Companies
V	How do you approach the role of internal audits to detect AI-based ethical consequences?	Both
VI	What do you think about the independency and the objectivity in the internal audits? From this perspective, how do you approach the role of external audits to detect AI-based ethical consequences?	Both

Table A.4: Questions related to Research Question II.

Number	**Questions**	**Applicable For**
I	How can the most critical ethical issues be solved with auditing firms regarding AI in the financial industry/banks?	Both
II	How could banks recognise their capabilities and the consistency of their strategies in order to decrease unintended ethical consequences of AI? How do they suffer from the complexity of AI? How can they develop those capabilities promptly?	Both
III	How do you think the data produced by AI applications can be explained?	Both
IV	How can financial companies solve the high data unavailability in AI technology and increase the transparency with auditing services?	Auditing Companies
V	How the financial services companies can protect and control their data and their customer's data? Can you explain it more regarding GDPR?	Both
VI	What do you suggest for algorithmic audits?	Auditing Companies
VII	To detect fraud, which one is more efficient, AI or Benford's Law? What do you think about an obligation from the algorithmic auditing perspective in order to control banks' AI detection mechanisms?	Auditing Companies

Table A.5: Questions related to Research Question III.

Part I: The New White Paper on AI		
Number	**Questions**	**Applicable For**
I	The European Commission wants to establish new binding requirements for the development and use of "high-risk" AI applications. As you may know, European Commission released a new white paper on AI, A European Approach to excellence and trust. What is your opinion on this invitation?	Both
II	What do you think about a voluntary labelling for no-high risk AI applications? How could it be applied for AI in auditing/banking?	
III	The Commission also indicated the limitations of scope of existing EU legislation. How can the scope be extended?	
IV	How can regulators build the readability of the AI algorithms?	
V	What could you suggest to European Commission to solve the gap in terms of AI and ethics in the legislation part? What else could be improved in this topic?	
Part II: Ethics Guidelines for Trustworthy AI		
Number	**Questions**	**Applicable For**
I	On 8 April 2019, the High-Level Expert Group on AI introduced Ethics Guidelines for Trustworthy Artificial Intelligence. As stated in the Guidelines, there are two elements of Trustworthy AI, namely an ethical purpose and technically robustness and reliability (The High-Level Expert Group on AI (AI HLEG), 2019a). What do you think about the EU framework "Trustworthiness"?	Both
Part III: Application		
Number	**Questions**	**Applicable For**
I	What do you think about the application of the regulations? Is it comprehensive/difficult/easy to apply?	Both

Appendix B

Categorisation of Participant Responses

In this Appendix, the categorised versions of the responses to the interviews are presented. During the categorisation, the data was managed by distinguishing different categories from the participants' responses. The categorisation was made considering the research questions of the thesis and the theoretical saturation. There are four main parts; namely to decrease unintended ethical consequences of AI, auditing services with ethical issues in AI applications, solutions to ethical issues on AI together with auditing services; and suggestions for AI legislations in the EU. The second-order categories are given in the following tables, in which the categories are alphabetically sorted.

Table B.1: Quotes related to second-order categories on unintended ethical consequences of AI.

Category	Quote
To decrease unintended ethical consequences of AI.	
A change in the strategy or the capabilities of the institutions is needed.	A: We, as a company, we have own risk management approach. We have certain sectors that we don't want to audit so we stay away from certain things. C: From a technical point of view, you need to realise existing techniques to make your model work...A lot of people don't know yet all the explainable sides of these deep learning techniques over these black boxes. It's sort of a fun way to remember it because in a cocktail you put lime and ice. These are two common techniques that will open up. E-I: We have an enterprise risk management framework which is already taking into consideration ethical risks which is linked to usage of AI. It would require a change in the strategy or the capabilities of the institutions. I also see a need to embed existing framework into day-to-day practices.
A mutualisation perspective is needed.	A: We also know that our shareholders that are requesting for certain return on certain dividends so they will have a piece of the cake. As an organising insurance organisation, you're also going to be constantly on the balance between providing coverage and also bad risks. In any company, shareholders expect that you would be generating a certain profit for them. E-II: Banks uses my data to make profit. What extent I would accept that my bank uses my data? Knowing that I can benefit from it, so it's for as long as it's a win win. H: Put your amount in a bank account and use it. It is good to know a lot of customer details but at the end, you still need to mutualise the risk. Otherwise we're not in insurance company anymore and so the difficult question is to find the right balance between good prediction at an individual level and good knowledge at the same time.
Adaptation is needed.	C: We are adaptive, our brain adapts, normally the strongest people adapt and know how they survive. AI models by default don't adapt so you need to have an operational monitoring tool to see it. You need to make sure that your model becomes adaptive.
AI is more powerful than Benford's Law.	A: AI is more powerful than Benford's Law. If you have data, if you have numbers together with the time, the stamp, the name or the location. Then AI becomes more powerful because you can take it bigger datasets.

continues in the next page...

AI is the solution, not the problem.	A: We can use algorithms in order to identify which transactions are more likely or suspectable than other ones. For a banking insurance company, they go with all the digital transformation things for reducing the cost base and increasing the speed in order to serve their clients. They still need to define what is the problem they want to solve. D: We have a lot of very powerful AI algorithms that can do beautiful tasks, not being very transparent but still being extremely useful. E-II: The purpose there is to automate things to increase productivity. They can invest in more added value stuff so obviously if you speak about productivity, you speak about cost savings. H: AI does not impact problems, it solves problems in the way you need to change your mindset and the way you interpret the results, so with AI it's possible to reduce the time to make decisions towards the customers to have a better price and also to have a better view on our obligations we have reserved.
An ethical policy is needed.	A: If you look at ethics, for example National Bank of Belgium, they already have guidelines on ethics, 10 - 20 years old. These guidelines are there so that's not changing, the only thing that's changing is that you as a company, you need to have an ethical policy and you need to define your code of conduct where you stand for so all of that. It's a company culture that we've created based on the tone at the top.
Anticipate the data.	E-II: I would say that rather than tackling the problem, after having created the insights we tried to actually anticipate it, so we tried to find out earlier to avoid problems later on.
Be aware of the risks of technology, be not too dependent on it.	G: A chatbot is like an interface and you have to connect to a secure system. So there are challenges coming here. I'm an entrepreneur, I love innovation, I love technology but more and more I am also aware of the risk of technology and I'm not even talking only about ethical risk but I am just thinking about how dependent we are becoming.

continues in the next page...

Be careful on what is happening behind the scenes in your tools.	A: Certain man and woman living together in the same household with the same type of income with the same type of expenditure, but woman get lower credit limit than men. There was an awkward example of how not having the right control mechanisms in place to detect this type of bias correlation. B: You have to take into account that when you have a system that can be spoofed quite easily, and you have to detect it and make sure that you don't get it. F: For NLP, it can occur in issues because our vocabulary is naturally gendered. If you're not careful on that, you're not aware on what's going on essentially behind the scenes, you may have problems with gender biased.
Collaboration with other departments.	B: We have a very good focus on digitalisation and innovation in compared with other banks. I think the innovation is really elaborated because almost everybody is involved in our bank and it's also linked to our environment and our culture. We put a lot of efforts in targeting everybody because we really use and believe in the wisdom of the crowd. D: We're doing models with all best intentions and with all the best practices in collaboration with people from IT to make sure that what we design robust applications. H: Once we're building a model, we have not only data scientist but we have also people from business. We have also actuaries and we have people from legal department.
Consider long term effects.	A: If you, as an organisation, are not ethical or if you don't have sufficient controls in place, then you could have a big issue in the future which will have a big reputation damage. It will also cost you a lot of money, so I think these are ethical and perhaps environmental and sustainable type of questions. I think they need to take them into account. On the short run they can still generate a lot of profit, but in the long run if they want to be there for 20-30 years, they need to start thinking and taking this into account. B: You have a lot of stakeholders internally, but you have also stakeholders externally. You also have to see this from a society perspective whether it is acceptable or not? C: You also need to understand whether your AI is going to impact a single person or a group of people. Is it going to have a short-term impact or long-term impact? D: Now, we are trying to protect again the privacy but even more also to understand how AI can have an impact on the citizens, what the impact of it on our clients, employers, employees and society will be.

continues in the next page...

Counterbalance the current bias.	A: If you look at the newspapers on the insurance, people will need to pay premiums because they're getting a lot of money from the insurance companies. If you're unethical at that side, you will suffer. It is an interesting balance that they need to strike with AI. E-I: We need to make sure that risk is properly mitigated or accepted by management. E-II: That's a very crucial point. You need to have data scientists that actually know how to handle it. They need to have representative, unbiased and balanced data samples. Otherwise you come up with some stereotypes. Unless you pay attention to that, you can quickly fall into discrimination which is not ethical. G: One of the bias of the current situation is that for example the north of the country has lower rates then the south of the country because Flanders is richer. Even if they put a common system, it may not work because this is the kind of thing to counterbalance some bias. If we want to make a fair system, we have to really think about all the bias.
Create a library of AI solutions for readability.	C: You have to recheck that rule from time to time. Creating a library of decision things can be a good thing so it's like central list of decisions that have been programmed. Everybody can read which decisions are made.
Data quality is important.	B: You have to think about the quality of the data that you have and you have to be able to challenge that quality. Then you should make sure that there is no biases at all and this is difficult one. Because it's all black box thing and you really don't know what is happening. E-I: The data shouldn't be abused. The data at the maximum should be aggregated and anonmyzed. G: This was the biggest challenge in bank was finding the data and making sure it's quality. The question around data quality and governance are fundamental. It's a huge challenge in any financial institution because they have billions and billions of data.
Data scientists can tell you the accuracy of algorithm, not ethics.	C: Data scientists tell you that your algorithm is accurate and you are just going to trust. Then it's going to decrease your cost so that's a concern for you. They don't care ethics because they don't understand it.

continues in the next page...

Define what unethical and ethical means to you.	A: Sometimes we hear from our clients "Help me detect unethical things in my data". The model itself is usually not discriminating, it's just doing what is learned out of the data. I think the big problem is that if you can define what unethical means to you. You need to work really hard with your organisation on defining what you believe is not ethical because it will be different for every organisation. C: What's ethical for you is not ethical for me. Ethics means the rights, it means something different to everybody. There's no golden rule for ethics something that's ethical here in Belgium is not ethical in US or in China and that's the difficult part.
Fix the biases.	B: In the Dutch language, there is an issue today in the chatbots because it doesn't work. I think there are less people speaking that you have a lot of dynamics here and there is an issue. C: Where the problems are for banks and insurance is that most AI systems are on the market and they will rely on old data and they will be cognitive biased. D: The intrusive data needs to be removed out of the model because it's not necessary for the purpose that we are trying to do. G: If we could keep on using this, it will just reproduce the bias overtime which is totally unfair of course. It can be really useful to fix the bias without changing the judicial intelligence of the past. So basically, all the decisions through all the tribunals in the country, you can probably have a much better understanding of what judgement you should give, if you can correct the bias of the past of course.
Hard to work with data scientists because of the different views.	B: The difficulty from my perspective is that people who are involved with this mostly data scientists. They're quite mathematical from thinking perspective and they are quite used to binary thinking 1-0. When you compare it to ethics, ethics is never black and whites. It's always grey. This is difficult to work with data scientists because of this totally different approach.
Keep a balance between explainability, transparency and interpretability.	C: In the insurance companies, you're going to use pictures from an accident to determine that somebody isn't fooled or not. These decisions are made much more accurate but they're less explainable. You have to balance all these two lines. H: Explainability and interpretability are two different things. You don't need to explain everything. At least, you should partly interpret. So for me, explainability is not mandatory. If I want to automate the reading of an email, I don't need to explain it. Of course, if I ask to make a price or customer, I'm able to explain why I come up with this.

continues in the next page...

Keep the human in the loop.	A: From the ethical sides, the last person should take the decision and evaluate the outcome of model. Chatbot supervisors are still humans. If you're a serious bank, you have somebody doing the monitoring this type of sentiment detection. B: I think it's also important that you make sure that there is a human control of it. C: AI is not something about data and algorithms but it's much more also about people and process. AI is just a system that can work in certainty and it gives you a number. The potential for fraud is 0.97 at the end, that's not the decision. It's just an insight. Somebody at the bank will make a decision. My point is that you can't blame AI for being not ethical. The risky part of chatbots is reinforcement learning technique which is dangerous. So you should always have a human in the loop to control it. D: We are using it chatbots as a decision support system for a human being who can use the AI algorithm as a recommendation and then still has the final goal to decide upon it. If it's becoming too sensitive, human being should still be in the loop to enhance oversee and correct overall the system. F: We have symptoms which don't grant loans systems because of the risk. So although the auto decision mechanisms can help to get initial markers down, in the end, person take the responsibility for that. H: You need to take into account ethics but you still need to keep the human in the loop at this point of time. We're not mature enough to have complete an AI ecosystem with no action of human.
Make a honest comparison between human being and AI system.	D: We need to have an honest comparison between human being versus automated system. AI systems which are often trained based on historical decisions of human beings so in the end, they will just copy the behaviour, hence make the same mistakes. We are also making decisions that are not easy to interpret applications. For example, classifying pictures are very difficult to understand. It's very difficult to understand why he decides that a picture of a cow. It is also very difficult for a human being to explain. H: Of course, AI makes mistakes but they will make mistakes less than the mistakes of humans.
Reverse the outcomes.	B: In the future, we will try to make it a little bit broader but what we do is checking the outcomes. For example, you have a model that uses data from females and males and you have a certain outcome of it and then we reverse all women into men and vice versa. From fairness aspects, we are checking the outcomes whether they are same or not.

continues in the next page...

Stakeholder analysis	B: When you partner with a third party for digital platforms or that kind of services, I think, you really have to do your homework very well to make sure that how robust they are. F: Understand who the stakeholders are and what could this possibly do to any of these stakeholders in digital platforms.
The risk depends on how you apply AI.	C: You can make your decisions explainable depending on which AI technique you've been using. If you're using just simple decision tree which you can hardly classify AI, it is much more transparent but not very accurate. D: I think it's how you apply it based on where the risk is. That's the same thing as an all technological revolutions, we make dynamite to help the coal mines but unfortunately the same dynamite was used to kill people. Is there something wrong with dynamite? No, we can use it in the construction industry and the mining industry but application of it sometimes can be wrong. Same thing applies here to AI, so there's nothing wrong with chatbots, NLP and in general AI by itself.
Train people.	E-I: The control functions themselves are not equipped in order to assess AI. Usually in control functions, you have people coming from the business. It needs very significant training of those people to make sure that it will be well informed and enlightened to control and supervise the way AI is being used by the defence. F: Training is very important for people working on these projects or on these systems. They are aware of ethical problems of risks. So we educate people.
Trust is important especially in finance and healthcare.	F: For us, finances and health are too very viral things to secure the stable life. So, the trust of the consumer in those sectors is essential. If you lose the trust of finance sector, you will lose the trust of your customers. Recovering that trust is impossible.

continues in the next page...

Understanding is important (AI, language, biases, etc.).	A: You need to have everybody in the organisation to understand what they're doing in the transformation...You need to define what bias you have, so you need to make sure that your model keeps on performing well. B: When we have a question of building a model, the first thing that we are going to do is to check what is the aim of this model and what is the possible damage based on the aim of the model. If the damage is non existing, we do not look at it anymore. C: Creating ethical AI in banking and insurance is a need for understanding the techniques but also it needs a multidisciplinary team understanding where you're gonna automate decisions based on AI. D: I think there is a lot of first made honestly by people who don't understand what AI is...People who don't understand what it is, they are exaggerating enormously. So for what we're doing with the bank with advanced analytics and AI, everybody is there and they will understand what the consequences are and will decide on very informed basis what the impact will be. Potentially, they will say yes or no based on that impact but it's not as black box as for everybody thinks it is. E-I: Every outcome that is produced by AI has to be in a position to be retrofitted to the original data and therefore we need to understand clearly decision mechanisms so that we can have full traceability between the outcome and the source.
Usage of the customer data dilemma.	A: If you're drunk, you won't get reimbursed about damage. As an insurance company, if you looked it up on Facebook which is private information, are you going to use it to grant somebody? This is the type of dilemmas for insurance companies but you can't just use it without your client knowing that you're using it. B: For using data, how far can you go in that and run? That is never black and white elements. It will become dark grey and then you have to stop it. It's very difficult to make this boundary. D: Banks are heavily using AI tools for those purposes (digital platforms) that you mentioned. You need to very carefully see what you will do with AI and what you will do with the data of a client. E-I: The main challenge is structuring the data in order to make a proper usage of them including the developing AI in order to get the benefits that we want. Can we use the client data, can we resell them? The real question that we have been handling for the moment is about ownership of the data. E-II: Asking the insurance company, you as a customer, "When do you believe what is my selected lifetime according to your model?" and they will tell you "five years" so the very nature of calculating your life expectancy by itself is something that comes with ethical concerns but also the elements that the company uses to determine that like the lifetime expectancy can also bring ethical issues.

Table B.2: Quotes related to second-order categories on the relationship between auditing services and ethical issues in AI applications.

Category	Quote
Relationship between auditing services and ethical issues in AI applications.	
Auditors can not solve the explainability, companies should do it.	C: It's not the audit that will explain your model, so the audit will see that it's not explainable and nothing was on that explainability side. The company will solve itself.
Auditors can not solve the problem.	B: I don't know even if the auditing companies are the most suitable one for looking at ethical questions for that. One of the ideas that we have, and we sometimes use but also in other elements, we have an external ethical board. An external advice is not obligatory to follow but it's still an advice. H: Audit will show the potential gap, lack and the problems then they could give us advice and sometimes respect over company so we could put some best practices and to put them in place. But audit itself I don't think they would solve the problem.
Consider the conflict of interest.	B: You have to make sure that you cannot outsource key banking issues. You have always an inherent conflict of interest. So there you have to be careful but this will not mean that you can never use an audit firm to support you in a consultancy solution but you have to be careful you choose.
Internal audits can be slow in AI and ethics audits.	F: A few of the banks set up actual ethics departments and kind of having overseeing view of everything but those are much slower and the ones, that we had come in contact with, are that are struggling.
Pro-active approach with external auditing services.	C: I would suggest not to do it internally but asking an external person. External auditors will see the issue easier. F: If you have one person looking at it, from their perspective having the audits you can get a clear view of where the big points are. If you are aware the risk points prior to any type of situation around risks, you will be prepared. So it is shifting pro-active rather than re-active in risk problems.

continues in the next page...

Put in place some user-friendly tool rather than internal or external audits.	H: To detect AI-based ethical consequences, we are making a model in fraud detection or pure automation. We're still trying to put in place some user-friendly tool with good governance. Governance has to be put in place. That is something that takes time, so the model it's something to take into account and that is what we are doing, we have a governance tool. We put the definition, we recover data.
Softness of ethics affect the audits.	A: If there is an impact in a financial statement do something with it, companies go to the audit committee and they report that the type of items. Same thing goes for ethics but it's very soft so it's not a clear like hard requirement.

Table B.3: Quotes related to second-order categories on the solutions of ethical issues in AI applications together with auditing services.

Category	Quote
Solving ethical issues in AI applications with auditing services.	
Auditors should also understand the different insights.	C: Process you have to understand every single part and you have to make sure that you have access internally to the right people who can give you the answers because you will need to have the different insights. E-II: Audits would have amended to make sure that what we do with the insights.
Audits should be done fairly.	E-II: Audit should be independent obviously. H: All audit is good as long as it is done in the fairway, in the collaborative way.
Auditors should understand what AI is.	E-II: Auditors should have also knowledge about AI. They should be aware that they should be able to understand what AI is about data, and what is the sensitivity of data, how may be putting together data which in isolation or not sensitive. G: They can look at the purpose of the algorithm. Why they do this? What is expected result? H: Professionals and people should understand what AI really means.
External auditors should gain the view on the internal corporate policies of the clients.	B: I think auditing should be linked to the company values. It has to be tailor made. E-II: In my opinion, the audit should also have a view on what are the internal corporate policies of their clients because there's a second side of the coin.
Having own challenger models in auditing services.	A: We have our own challenger models and we run the same data through our challenge almost to see if we have similar results. That's something that is common in the insurance sector for valuation to derivative valuation, so we have our own right pricing models to check the pricing at the clients and to make sure that it's in line with what we expect.

continues in the next page...

Standards are needed for external audits.	A: You need to define what exactly your rules are. As an audit firm, you need to go to rules and standards. C: It's an absolute important thing to get guarantee label for AI. For most decisions that are having an impact on our lives such as medical things and food, we have a guarantee label. The same thing should happen to AI systems. E-I: For example, there is an association of internal auditors, it is producing standards to support auditors in doing their job. Having similar standards to audit AI is certainly something that would be useful for internal audit or another lines of defense.
Test the tool, think the whole process.	A: If you are an internal auditor, the big question you always need to answer, is my tool working? You need to even test your tool that it's really working to do an audit because the worst thing you can do is to use a tool that is not catching any risks that you want to detect. Secondly, if you're using that tool like a design tree to find something, in audit, one of biggest thing is always you need to come up with an audit approach. So you always need to go through the whole process of thinking "What is the approach for me about the risks I want to cover?". D: I think it's more important to understand and to be very well focused on testing the capabilities of a model first.
Try to see hidden purposes in the algorithms.	G: When you have an algorithm, trying to see if there's not a hidden purpose that would be the main thing to audit.
Understanding each other in multidisciplinary teams.	C: We all use the word "trust" but we all mean something differently. That's why, also in audits, when even if you make an effort to do for multidisciplinary team you have to make sure people understand each other.

Table B.4: Quotes related to second-order categories on AI legislations in EU.

Category	Quote
Suggestions for AI legislations in the EU.	
Application of regulations is difficult to apply.	B: It's very difficult to cope with it. As a larger bank, it is still OK but if you're a quite small company, it is really impossible. C: World is very complex and it's impossible to put your complexity into the system. Look at COVID-19, you have the best scientists worldwide looking at it and nobody gets it right at the moment because world and the nature is more complex than what we have. There's a lot of things even in finance that are very subjective. Sometimes with the best algorithms, you can't predict so it's very difficult. D: Where I personally have a problem with is the whole discussion about bias of AI and especially biased corrections of AI that we need to have. The legislation is a little bit go into that direction that we should correct biases. I think it will become a very difficult thing...Difficulty is removing data that you don't know where it will end up. E-I: It's very difficult because of taking into consideration financial constraints and technical constraints so it could be an additional layer of legislation. F: From an ethical aspect, it is great to list these principles of EU framework. It was good for step because you need a kind of code. What would you do after that? How do they be applied? These are the issues.
Clear guidelines, rules, and principles are needed.	A: There are lots of frameworks in the market. This will keep on evolving, and it's important to follow that up. I think everything is on the table now but it's really waiting the European Union as a power to really say "These are the principles..."and if there is a standard clear rules, I think auditors can do something with that and they can check the rules against the application. So that is clearly on the strategic piece. There's one central database of the government that will help the whole sector and that's clearly a point to improve the control function. D: We need clear guidelines and rules about what we can do and what we cannot do. E-I: In order to add value to the business, benchmarks are made available to help business to help auditors etc. but it needs to be more concrete. Also they should not leave these principles as too much room for interpretation.

continues in the next page...

Consider certain values.	A: In European Union, we have also certain values. European Union has all set of values and these values are different from China for example. That's why in China it's possible to use face detection by governments to do whatever they want to do but in EU we have problems with that because it goes against our values. We should really set policy in order to be compliant with these values. B: On the other hand, we have a global perspective even large countries already working a lot on AI and taking into account ethical perspectives from the way the culture. For example, China use face recognition, from European perspective this is not OK...also it's like a culture and how you can embed this and I think you should also try to get this responsible behaviour view on things.
Consider the co-ordination and globalisation.	B: You have the globalisation so you really have to take into account it...you also see now for example in GDPR, it is becoming more and more important so in other parts of the world. You see also in the US. E-I: It will be very important to liaise with other like the US, India or China in order to get as much alignment as possible.
Do not kill the innovation.	E-I: We have to put in place regulatory framework that would avoid the abuse of consumers data for AI but at the same time we have to be careful not to develop legislations that would be considered as being counterproductive versus the appropriate of innovation. G: In Europe, we are already late but in other countries such as China and the US, they have tested everything because they were allowed to, the ethical privacy, dependency, consent and everything but also not killing innovation because we live in an interconnected world. H: We still need to be flexible so that data scientist can work on models so it should not be restricted too much. If you put regulation to district everyone, I don't think that it's the goal. So this is important but it has to be done in the smart and interest away.
EC should reflect their purpose.	B: Sometimes they want to regulate something which is not mature enough which means that regulation is totally outrageous for the purpose and you get much too much regulations and it doesn't fit any longer the aim that initially helped from a protection perspective. G: Europe should also reflect their own purpose on this works. If their purpose is just to annoy people, everyone will reject them but if they want to be seen as just an organism that helps everyone to work better and fairer, then that's different.

continues in the next page...

Follow data science standards.	H: I think the key will be to follow the standard of data science. If it is not a data scientist playing with the tool, it means that it is someone who doesn't have the knowledge and hasn't been trained correctly but will use the tool.
Hard to define risks.	C: You're going to have a chain of applications and maybe your application is not high risk but it's linked to a higher risk. It's because of an error in your application... They look at single applications. It's like in a plane, every individual component might be perfect but in airline crashes and what you see, it's because the threshold of one application was good enough to pass information to the next application where it was good enough to pass the next application and the final application fails but it's because of the very early one. I find it risky to classify it like this because the world is a little bit more complex. D: What is the high, medium or low risk. Otherwise, everybody starts applying the rules based on their interpretability then it doesn't make sense to do it. We need to look at what can go wrong as a risk management. Are we willing to accept that? Will you tell what high medium or low risk is? So it needs to be an informed judgment. Even if there is a judgment, it should be an informed judgment. E-II: You need to have a proper transparent description of the criteria used to label an application as medium, high or low risk based either on the data or the input data. F: My first question to them would be how do you define high risk, one of the things with AI is you can create an AI application and use it in a low risk situation but as soon as it's taken out of that context it would be considered high risk so how do you define that in the first place? I think, yes, it makes sense to have a bit stricter rules in terms of high risk. G: If you ask people to voluntarily label something, I think the huge majority will do it properly at least as long as they can. However, sometimes labelling will be challenging what is a low risk, a medium risk, or a high risk.
Having more balanced view is important.	D: Where you find good balance, being not over critical to AI applications as to which all other decision systems that we have. IT-based or human based decision making, the same level of control with the same level of objectivity needs to be maintained. E-II: At least you have some transparency and the collaborative work around the topic and having the view of multiple people allows you to have a more balanced view and, hopefully have more balanced constraints in the end.

continues in the next page...

Labelling can be mandatory.	E-II: It would actually force people into thinking about in the very early stages of development. So that forces you to take into consideration aspects. If such a label becomes mandatory, you get people to think about it.
Problem is the lack of accountability.	B: When it's all in in house, it's OK. But when the AI elements are outsourced where do you put the responsibility and accountability? How will you cope with? C: Data scientist does not know where his application is going to be involved. If they don't tell me where my application is going to be used, I cannot inform them about the dangers and it's definitely not the data scientists. Who is accountable for this? F: The problem is due to the lack of accountability. You have to decide who is going to take the responsibility for it?...The problem is no one is willing to take responsibility.
Regulators should guide and audit.	F: You can read and understand the terms but applying is another thing. Maybe EU needs to regulate it. G: You can guide or audit companies just to make sure that whether their labelling is correct or not.
Standards and boundaries are needed for the voluntary labelling.	F: People will only go to that line if it's low risk but if you have this voluntary labelling system like you have to meet these five conditions and then you can label system as transparent...in a low risk context and then added risk and have it actually risk hidden. Boundaries are important for it and extremely difficult. H: We can say that it's a low level for explainability but it is a high level for fairness or away so I think it should be allowed because you will not be able to add an effective regulation globally for all models depends in base level will be a huge points.
The difficulty is the translation of ethics.	A: These are principles which are extremely different to implement because you need to translate your principle to make it practical to rule. B: It is always difficult especially when you have to translate it into working guidelines for people who have to work with it on a daily basis. D: I sometimes ask myself how it will be possible to translate certain regulations on AI but especially on ethics, because translating ethics in general into rules is a very difficult thing.

continues in the next page...

The white paper is so narrow.	B: I hope they will broaden the existing legislation a little bit. C: It's just focusing on one end application for example a risk application or fraud application. The world is where you're going to have two autonomous vehicles meeting each other. For example, if they have an accident, you basically have two high risk applications competing with each other. The white paper doesn't say anything about the interaction between two high risk applications so it tends to take a human versus a system whereas we're going more and more into a world where we have a system against the system and it's nothing mentioned there about the trustworthy part.
Understand the language, the problems, and the purposes of AI.	A: The scope always starts with define what is AI and what is not AI. D: I think it's more important to understand and to be very well focused on testing the capabilities of a model first. How well does it work and then if necessary you can also focus on how does it come to this result, very well understanding the language of AI. It's with those people you need to make a judgment on it. E-II: We interacted iteratively and so we eventually had a model that was able to predict something like 85% of the documents with accuracy. So that was understanding written languages. F: I think regulators should educate themselves first and understand what the problems are. Because AI and ethics are very complicated. G: I'm sure through privacy some of those questions are already coming but maybe it doesn't cover enough of the purpose. For the regulators, I would focus on the purpose of why we collect data and why we build models.
Voluntary labelling can be useful for establishing the trust between parties.	A: In Belgium, you can outsource the activities of calculating the payroll. The problem is a company that you don't see what's happening in the centre. So what the solution there is that they are audited every year or multiple times a year based on a list of criteria and then they issue reports to your company so you know that they are taking care of my payroll but I have a report that everything is really functioning like it should be on there. And so that gives them trust between these two parties. That could help you in establishing relationships between parties in the market. I think voluntary labelling would help and I think we've already seen this type of certifications in other spaces like payroll services or more.

continues in the next page...

Voluntary labelling will not work in practice.	B: Voluntary labelling will be difficult when other places in the world are not doing this and if this should mean that it will cost much more to set this up for company and I think it will not work. E-I: Financial institutions lie to be bound by too many norms. And if it is being left to this institutions I'm afraid it won't work because institutions will always find a reason to criticise one another. It is a nice ambition but not working in practice so I think that if it is the objective to have that kind of labelling, they should put in a place of binding legal requirement.

References

Accenture. (2017). *Technology Vision for Insurance* (Tech. Rep.). Retrieved from https://www.accenture.com/t20170418t020959{_}{_}w{_}{_}/ph-en/{_}acnmedia/accenture/conversion-assets/nonsecureclients/documents/pdf/2/accenture-technologyvision-insurance-2017.pdf

Berscheid, J., & Roewer-Despres, F. (2019). Beyond transparency. *AI Matters*, *5*(2), 13–22. doi: 10.1145/3340470.3340476

Buchanan, B. G. (2019). *Artificial Intelligence in Finance* (Tech. Rep.). The Alan Turing Institute. Retrieved from https://www.turing.ac.uk/research/publications/artificial-intelligence-finance doi: 10.5281/zenodo.2612537

Chan, D. Y., & Vasarhelyi, M. A. (2011). Innovation and practice of continuous auditing. *International Journal of Accounting Information Systems*, *12*(2), 152–160. Retrieved from http://dx.doi.org/10.1016/j.accinf.2011.01.001 doi: 10.1016/j.accinf.2011.01.001

Christian, S., Kevin, H., Karrie, K., & Cedric, L. (2014). Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms Christian. In *Proceedings of the data and discrimination: Converting critical concerns into productive inquiry preconference.* Seattle. doi: 10.1145/3375627.3375852

Cramer, H., Jean, G.-G., Springer, A., & Reddy, S. (2018). Assessing and Addressing Algorithmic Bias in Practice. *Interactions*, *25*(6), 58 – 63. Retrieved from https://doi.org/10.1145/3278156

European Commission. (2019). *A definition of Artificial Intelligence: main capabilities and scientific disciplines.* Retrieved 2020-05-09, from https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

European Commission. (2020). *White Paper on Artificial Intelligence A European approach to excellence and trust.*

Financial Stability Board. (2017). *Artificial Intelligence and Machine Learning in Financial Services - Market Developments and Financial Stability Implications* (Tech. Rep.). Financial Stability Board. Retrieved from http://www.fsb.org/2017/11/artificial-intelligence-and-machine-learning-in-financial-service/

Fourie, L., & Bennett, T. (2019). Super intelligent financial services. *Journal of Payments Strategy & Systems*, *13*(2), 151–164. Retrieved from https://www.ingentaconnect.com/content/hsp/jpss/2019/00000013/00000002/art00008

Fruhan, W. (1979). *Financial Strategy.* Homewood: Irwin, R.D.

Fuscaldo, D. (2020). *ZestFinance Using AI To Bring Fairness To Mortgage Lending.* Retrieved 2020-01-29, from https://www.forbes.com/sites/donnafuscaldo/2019/03/19/zestfinance-using-ai-to-bring-fairness-to-mortgage-lending/{#}29be69427f2d

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* London, England: Aldine Transaction.

Gossett, S. (2020). *10 AI in Banking Examples You Should Know.* Retrieved 2020-01-28, from https://builtin.com/artificial-intelligence/ai-in-banking

Green, B., & Chen, Y. (2019). Disparate interactions: An algorithm-in-the-loop analysis of fairness in risk assessments. In *Fat* 2019 - proceedings of the 2019 conference on fairness, accountability, and transparency* (pp. 90–99). Atlanta. doi: 10.1145/3287560.3287563

Hall, S. (2017). How Artificial Intelligence is Changing the Insurance Industry. *CIPR Newsletter*(August).

IIA. (2013). *Performance Standard 2050: Coordination, The International Standards for the Professional Practice of Internal Auditing* (Tech. Rep.). Institute of Internal Auditors Research Foundation, Altamonte Springs, FL.

Ionescu, L. (2017). Errors and Fraud in Accounting: The Role of External Audit in Fighting Corruption. *Annals of Spiru Haret University. Economic Series*(4), 29–36.

Kahn, J. (2020). *The problem with the EU' s A.I. strategy.* Retrieved 2020-02-29, from `https://fortune.com/2020/02/25/eu-a-i-whitepaper-eye-on-a-i/`

Larsson, S., Anneroth, M., Felländer, A., Felländer-Tsai, L., Heintz, F., & Cedering Ångström, R. (2019). *Sustainable AI - An inventory of the state of knowledge of ethical, social, and legal challenges related to artificial intelligence* (Tech. Rep.). AI Sustainability Center. Retrieved from `https://portal.research.lu.se/portal/en/publications/sustainable-ai(e2fa1b6a-860e-44b0-a359-fbd842c363db).html`

Li, G., Deng, X., Gao, Z., & Chen, F. (2019). Analysis on ethical problems of artificial intelligence technology. In *Acm international conference proceeding series* (pp. 101–105). Nanjing. doi: 10.1145/3341042.3341057

MIT SMR Connections. (2020). *How AI Changes the Rules : New Imperatives for the Intelligent Organization* (Tech. Rep.). SAS.

Muñoz-Izquierdo, N., Camacho-Miñano, M. D. M., Segovia-Vargas, M. J., & Pascual-Ezama, D. (2019). Is the external audit report useful for bankruptcy prediction? Evidence using artificial intelligence. *International Journal of Financial Studies*, *7*(20). doi: 10.3390/ijfs7020020

Murphy, L., Huybrechts, J., & Lambrechts, F. (2019). The Origins and Development of Socioemotional Wealth Within Next-Generation Family Members: An Interpretive Grounded Theory Study. *Family Business Review*, *32*(4), 396–424. doi: 10.1177/0894486519890775

Pascual, A. (2015). *Future Proof Card Authorization.* Retrieved from `https://www.javelinstrategy.com/press-release/false-positive-card-declines-push-consumers-abandon-issuers-and-merchants.`

Pau, L. (1987). *Artificial Intelligence in Economics and Management* (II ed.). Amsterdam: Elsevier Science Publishers.

Qu, S. Q., & Dumay, J. (2011). The qualitative research interview. *Qualitative Research in Accounting and Management*, *8*(3), 238–264. doi: 10.1108/11766091111162070

Quah, J. T. S., & Chua, Y. W. (2019). Chatbot Assisted Marketing in Financial Service Industry. In L.-J. Zhang, A. Musaev, & J. Eduardo Ferreira (Eds.), *Proc. 16th international converence on services computing - scc 2019* (pp. 107–114). San Diego: Springer Nature Switzerland. Retrieved from `http://link.springer.com/10.1007/978-3-030-23554-3` doi: 10.1007/978-3-030-23554-3

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., ... Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Fat* 2020 - proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 33–44). Barcelona. doi: 10.1145/3351095.3372873

Rauch-Hindin, W. B. (1988). *A Guide To Commercial Artificial Intelligence : Fundamentals and Real-World Applications.* Prentice Hall.

Riikkinen, M., Saarijärvi, H., Sarlin, P., & Lähteenmäki, I. (2018). Using artificial intelligence to create value in insurance. *International Journal of Bank Marketing*, *36*(6), 1145–1168. doi: 10.1108/IJBM-01-2017-0015

Rohner, P., & Uhl, M. W. (2018). Robo-Advisors versus Traditional Investment Advisors: An Unequal Game. *The Journal of Wealth Management*, *Summer2018*, 44–51.

Russell, S., & Norvig, P. (2014). *Artificial Intelligence : A Modern Approach* (III ed.). Pearson Education Limited.

Shah, H. (2018). Algorithmic accountability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*. doi: http://doi.org/10.1098/rsta.2017.0362

Simon, A., Yaya, L. H. P., Karapetrovic, S., & Casadesús, M. (2014). An empirical analysis of the integration of internal and external management system audits. *Journal of Cleaner Production*, *66*, 499–506. Retrieved from http://dx.doi.org/10.1016/j.jclepro.2013.11.020 doi: 10.1016/j.jclepro.2013.11.020

Stamatis, D. H. (2003). Failure mode and effect analysis: FMEA from theory to execution. *ASQ Quality press.*.

Stanwick, P. A., & Stanwick, S. D. (2001). Cut Your Risks with Environmental Auditing. *Journal of Corporate Accounting & Finance*, *12*(4), 11–14. doi: 10.1002/jcaf.2403.abs

Suddaby, R. (2006). Determinants of product innovation in small firms: A comparison across industries. *Academy of Management Journal*, *49*(4), 633–642.

The High-Level Expert Group on AI (AI HLEG). (2019a). *Ethics Guidelines for Trustworthy AI.* Retrieved from https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines{#}Top

The High-Level Expert Group on AI (AI HLEG). (2019b). *Policy and investment recommendations for trustworthy Artificial Intelligence.* Retrieved from https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence

Vasarhelyi, M., & O'Leary, D. E. (2000). *Artificial Intelligence in Accounting and Auditing* (IV ed.). Princeton: Markus Wiener Publishers.

Vassiljeva, K., Tepljakov, A., Petlenkov, E., & Netsajev, E. (2017). Computational intelligence approach for estimation of vehicle insurance risk level. In *Proceedings of the international joint conference on neural networks* (Vol. 2017-May, pp. 4073–4078). Alaska. doi: 10.1109/IJCNN.2017.7966370

Wang, I. Z., & Fargher, N. (2017). The effects of tone at the top and coordination with external auditors on internal auditors' fraud risk assessments. *Accounting and Finance*, *57*(4), 1177–1202. doi: 10.1111/acfi.12191

Zarifis, A., Holland, C. P., & Milne, A. (2019). Evaluating the impact of AI on insurance: The four emerging AI- and data-driven business models. *Emerald Open Research*, *1*(15), 1–18. doi: 10.35241/emeraldopenres.13249.1