

MESG
MESTRADO EM ENGENHARIA
DE SERVIÇOS E GESTÃO

**People and (Ro)Bots: Trustworthy AI and Machine Learning
Applications for Services Automation**

Raquel Maria do Espírito Santo Faria

Master Thesis

Supervisor at FEUP: Prof. Gabriela Beirão

Supervisor at INESC TEC: Prof. Ana Torres



2023-07-20

This thesis work is dedicated to my love, David, who has been a constant source of support and encouragement during the challenges of school and life. I am truly thankful for having you in my life. This work is also dedicated to my parents, brother and sister, Carlos, Ivone, Daniel and Claudia, who have always loved me unconditionally and whose good examples have taught me to work hard for the things that I aspire to achieve.

Abstract

Artificial Intelligence (AI) has altered our world in an extensive number of ways. Although its application has its benefits, the underlying issues surrounding trust, privacy and security, biases and ethics, for example, need to be understood, not only by the organizations that want to implement and use it, but also by the users that are susceptible to its vulnerabilities.

To better understand the impact of these themes on users' perceptions and trustworthiness, and its applicability in a service automation context, this dissertation reviews the current literature on artificial intelligence, trustworthiness and privacy and security concepts, as well as conducts a qualitative study that analyses users' perceptions and views on the use of machine learning artifacts such as virtual assistants, chatbots, recommender systems and predictive analytics. Three data sources were used to uncover findings: reviews from an online source, and interviewees directed to ten general users and ten experts in the information technology field. Thematic analysis was used to uncover common themes from the data analysed.

Finally, the results obtained from the qualitative study were discussed, which highlighted the concerns users' face when utilising these machine learning artifacts. The personal safety and privacy category were recognized as the main risks, by all data sources, which emphasizes the concern on this topic, other themes such as biased algorithms and information were also referenced, which indicates that users have negative perceptions towards these AI tools. This study also brought to light benefits and challenges that organizations may encounter when implementing these machine learning applications, which allows organizations to better understand and make informed decisions upon integrating these tools. Additionally, this dissertation offers theoretical guidance for organizations that wish to implement these machine learning artifacts in a service automation context.

Acknowledgments

First and foremost, I express my heartfelt gratitude to Professor Ana Torres and Professor Gabriela Beirão, my thesis advisers, for their unwavering support, patience, and motivation throughout every stage of this academic endeavour. Their selfless sharing of expertise has been invaluable from the initial to the final phase.

I am immensely grateful to the University of Porto and FEUP (Faculty of Engineering of the University of Porto) for offering me a truly unique challenge and experience that pushed me beyond my comfort zone. Their warm welcome to Porto has made my academic journey all the more enriching and memorable. I would also like to acknowledge my professors and classmates for their guidance and support throughout my academic journey.

Last but not least, I want to thank my family for all their love and patience in making this journey meaningful.

Table of Contents

1	Introduction.....	10
1.1	Project background	10
1.2	Problem Description	10
1.3	Research Questions	11
1.4	Report outline.....	11
2	Literature Review	12
2.1	Artificial Intelligence and Machine Learning	12
2.2	Trustworthy AI.....	14
2.3	User Privacy and Security	16
2.4	Bibliometric Analysis of Artificial Intelligence	17
2.4.1.	Methodology	18
2.4.2	Results Analysis.....	20
3	Methodology.....	24
3.1	Data Source and Criteria for Selection	24
3.2	Data Consolidation and Extraction	25
3.3	Data Analysis	26
4	Results.....	28
4.1	Results from Google play store reviews	28
4.2	Results from the analysis of interviews	30
5	Discussion.....	35
6	Conclusion and Future research.....	43
6.1	Conclusion	43
6.2	Limitations and Future Research	44
	References	46
	APPENDIX A: Interview Guide	51
	APPENDIX B: Consent Form.....	52
	APPENDIX C: Codes from NVIVO	53
	APPENDIX D: Word Frequency Cloud for the four Google Play store applications ..	55

List of Tables

<i>Table 1. Users' trustiness perceptions of ML Artifacts</i>	<i>13</i>
<i>Table 2. Top keywords with the highest frequency.....</i>	<i>20</i>
<i>Table 3. Clusters' themes and sub-themes</i>	<i>22</i>
<i>Table 4. Data Source of ML Artifacts.....</i>	<i>24</i>
<i>Table 5. Demographic data of Users'.....</i>	<i>25</i>
<i>Table 6. Demographic data of Experts.....</i>	<i>25</i>
<i>Table 7. Concerns from Customer Reviews</i>	<i>28</i>
<i>Table 8. Overview of the concerns from the three sources</i>	<i>35</i>
<i>Table 9. Common concerns expressed by interviewees.....</i>	<i>36</i>
<i>Table 10. Concerns expressed solely by the expert group</i>	<i>37</i>
<i>Table 11. Benefits and Challenges for Org. based on User and Expert perspectives.....</i>	<i>39</i>
<i>Table 12. Strategy implementation for Organizations</i>	<i>40</i>

List of Figures

<i>Figure 1. Number of publications published in the period January 2010 - March 2023</i>	<i>19</i>
<i>Figure 2. Number of publications per country since 2010.....</i>	<i>19</i>
<i>Figure 3. Top research areas by number of publications</i>	<i>20</i>
<i>Figure 4. Network visualization of keywords</i>	<i>21</i>
<i>Figure 5. Density visualization of keywords</i>	<i>22</i>
<i>Figure 6. Benefits expressed by interviewees.....</i>	<i>30</i>
<i>Figure 7. Disadvantages expressed by interviewees.....</i>	<i>31</i>
<i>Figure 8. Concerns expressed by interviewees</i>	<i>33</i>

List of abbreviations

AI: Artificial Intelligence

EU: European Union

FCRA: Fair Credit Reporting Act

FERPA: Family Educational Right and Privacy Act

GDPR: General Data Protection Regulation

HIPAA: Health Information Portability and Accountability Act

INESC TEC: Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência;

IT: Information Technology

ML: Machine Learning

UK: United Kingdom

USA: United States of America

WoS: Web of Science

1 Introduction

1.1 Project background

Artificial Intelligence (AI) has transformed our world and has revolutionized the service industry. The automation of services has benefited organizations, by making them more efficient and making users experience personalized, efficient and reliable by reducing human errors significantly, resulting in an overall positive experience.

AI is a multidisciplinary field combining science and engineering aiming to create intelligent machines (Benaich & Hogarth, 2022). AI has seamlessly integrated into our daily lives and improved it immensely, in sectors such as healthcare and transportation and navigation, for example. As stated by (Benaich & Hogarth, 2022), everything that is encountered in our modern world is some form of a product of intelligence. And this could not be truer. Recently the rise of AI tools such as ChatGPT and other tools of this nature that can create, design and process information such as images, videos and others, without critical input, has revolutionized the way we interact. Whenever users engage with AI applications, it improves the way we interact and mostly provides us with accessibility to quick answers. Businesses are integrating these AI tools as a strategy to impulse efficiency in their processes (Dwivedi et al., 2021).

Companies report that AI fosters revenue in sales and marketing while reducing costs in supply chain management and manufacturing functions (Sadok, n.a). For companies, AI technologies entail several benefits: increased revenues, sales reinforcement, price segmentation from the customer database, customer satisfaction, lower product returns, and cost reduction (Syam & Sharma, 2018).

As companies can identify the many benefits, they are now making this shift, prioritizing technology over traditional processes and making it a recurring and common theme. A study conducted by experts, demonstrated that 95% of all customer interactions will be assisted by AI channels by 2025 and that chatbot applications will be the predominant choice (Anica-Popa et al., 2021). The benefits of the digital implementation are greater than ever before. While companies continuously choose to enforce these technologies due to the benefits associated, many users are highly relying and depending on these technologies.

1.2 Problem Description

AI and service automation can bring many opportunities, however it can also bring some challenges for businesses. In a more general approach referred by (Frank et al, 2019), one concern is the replacement of workers by AI, increasing concerns over unemployment and the overall function of the job industry. Also service automation may increase a lack of human interaction which can also pose as a challenge to organizations. Consumers who are not engaged in this new customer experience and have negative emotional states towards it, such as frustration, detachment, isolation or alienation may reject the utilization of this technology (Bolton et al, 2018). Also the use of AI applications in a critical field, such as medicine, can outline new risks. Biased diagnoses or treatments, AI robot autonomy and security and privacy concerns regarding medical information are just a few of the potential dangers of AI that can potentially harm a patient. For example, questions have arisen about whether or not AI can exercise the same rights as doctors (Sunarti et al., 2021). Therefore, AI entails several concerns for consumers. Biased AI algorithms, security and privacy concerning personal data and even

ethical concerns all diminish and have implications on user trust. It is essential to understand how to mitigate this risk and in what contexts they might emerge. As such, we need to know how customers interact with these applications and their trustworthiness, as they are becoming a big part of any company's service infrastructure. Understanding perceptions and concerns around the different AI applications in various fields is important.

The main challenge that organizations face, is finding the correct balance between utilizing these AI tools for their full potential and in an efficient manner, while respecting, developing and maintaining relationships with their customers, as this is the core of their value proposition. (Faria et al., forthcoming).

It is quite unknown how these AI tools actually impact services on different levels. Even though the utilization of these applications are widespread, there is a need to understand the ultimate benefits, challenges and concerns from a consumer perspective and on a general level. While many studies have concentrated on these tools in specific contexts (Bradlow et al., 2017; Nithya & Ilango, 2017), such as in a specific company context or solely focusing on one particular AI application, this research intends to uncover how consumers interact with these tools and their perceptions and levels of trust. As organizations are increasingly implementing AI in their business strategies, it is crucial for them to understand where consumers stand and how to efficiently implement AI without jeopardizing their business and their customer interests.

1.3 Research Questions

To answer this problem, this study focuses on responding the following research questions:

R1: What are user's trustiness perceptions of Machine Learning Artifacts?

R2: What are the benefits and the challenges organizations face, when implementing these technologies?

R3: What strategies enhance user trust, address concerns and ensure responsible and secure utilization of these technologies?

1.4 Report outline

This dissertation starts by giving a comprehensive analysis of the research literature (chapter 2) relating to topics of artificial intelligence, machine learning, trustworthy AI, user privacy and security and a bibliometric analysis to understand main privacy and security issues in academic research. Chapter 3, presents the main methods used in the dissertation in order to respond to the research questions. It indicates all the procedures and techniques to extract and analyse data. In chapter 4, it is presented the results that derive from the methodology chapter. Chapter 5 discusses the results encountered in chapter 4 and provides key findings and insights. Finally, chapter 6 presents the main conclusions and discusses limitations and future research opportunities.

2 Literature Review

2.1 Artificial Intelligence and Machine Learning

Artificial intelligence (AI) is a broad discipline aiming to create intelligent machines instead of the natural intelligence that humans and animals demonstrate. AI has become a global term that seizes the long-term ambition of the field to construct machines able to emulate and excel in the full range of human cognition (Benaich & Hogarth, 2022). Machine learning (ML) is a part of AI that usually uses statistical methods that let machines "learn" directly from data without explicit instructions on how to do it. This process is known as "training" a "model" using a learning "algorithm" that progressively improves model performance on a specific task (Benaich & Hogarth, 2022).

However, there is no consensus around the definition of Artificial Intelligence (AI), with various authors proposing several definitions over the years. One of the early AI definitions was introduced by (Bellman, 1978) as an automated process associated with human thought, including learning, problem-solving, and decision-making. According to Russell & Norvig (2010), AI is the study of agents that act after taking in environmental cues from their surroundings. (Dwivedi et al., 2021) defined AI as the increasing ability of machines to take over specific roles and tasks currently performed by humans in the workplace and society.

The different areas of AI have been divided in various ways over the past years. (Mccarthy, 2007) divided AI into twelve branches and topics: logical AI, search AI, pattern recognition, representation, inference, common sense knowledge and reasoning, learning from experience, planning, epistemology, ontology, heuristics, and genetic programming. The key point is that these systems use algorithms to gather, examine, compare, and predict as much information as possible. It was also noted that AI still has a long way to progress in reasoning. (Zadeh, 2008) states that the challenge of reasoning relies on two of humans' outstanding abilities. Firstly, the ability to reason, communicate and act rationally in the face of ambiguity, uncertainty, inadequate information, and partiality of truth. And secondly, the capacity to complete a wide range of physical and mental tasks without any calculations or measurements. Automating these abilities, particularly automating natural language understanding, is required to achieve machine intelligence on par with humans.

A more recent study by (Collins et al., 2021) specified a different range of AI branches from Expert Systems, Machine Learning, Robotics, Natural Language Processing, Machine Vision, and Speech Recognition. Expert Systems are defined as systems that try to represent knowledge and simulate human behavior when solving a problem. These authors conclude that machine learning is increasingly used and is a way to improve the method's results as it continuously gets more data. AI can also be applied to robotics, especially in manufacturing (also known as Industry 4.0). Natural Language Processing is a way of converting human language into a representation for computers to understand, and it is the base for AI speech recognition applications. Finally, machine vision is a way of analyzing images using an algorithmic inspection.

Machine Learning is a set of algorithms that enable systems to learn. Systems with machine learning capabilities may interact with data, learn from experience, and create intelligent solutions based on prior knowledge (Jarrahi, 2018). Several ways of applying machine learning to applications were found, some of which we use daily and might not even acknowledge. Based

on the literature it was identified four main types of machine learning applications: chatbots, virtual assistants, recommender systems, and predictive analytics.

AI implementation intrinsically depends on the amount and quality of personal information organizations gather from their customers. However, companies must have ethical considerations and choose the right balance between personalization advantages and privacy risks (Inman & Nikolova, 2017). To take advantage of an AI-powered shopping experience, customers are being increasingly pressured to share private information, which can easily make them mistrust AI and raise privacy issues. Although, privacy concerns regarding the collection and use of personal information for marketing purposes already exist, AI broad use is raising those concerns, as users' data is being collected and processed by machines and algorithms (Anica-Popa et al., 2021).

Table 1 shows the most relevant privacy and security challenges, barriers of use, and impacts on trust identified in the literature for the four machine learning types.

Table 1. Users' trustiness perceptions of ML Artifacts

ML Artifact	Features	Privacy and Security Challenges	Barriers of use	Impacts on Trust	Reference
Voice Based Digital Assistant/ Personal Assistant/ Smart assistant	Real-time observation and processing of complicated data;	Storage and use of data	The system uses passive listening	Those valuing this technology are ready to sacrifice privacy to reap its benefits.	(Vimalkumar et al., 2021); (de Barcelos Silva et al., 2020)
	Undertakes tasks on behalf of users;	Continuous analysis of user behavior	The user's personal information is not secure.	Privacy concerns limit the adoption of technology.	
	Personalization; Anthropomorphism;		Misunderstanding of the use and storage of data	The tendency for users to only do simple activities with the technology.	
Chatbots	The first line of support	Reduced access to the companies' personnel	Interpretation issues	The tendency for users to only do simple activities with the technology. E.g., users will not complete transactions	(Følstad et al., 2018); (Cheng & Jiang, 2020); (Hasal et al., 2021)
	Available 24/7		Misunderstanding of use and storage of data		
	No waiting or queue	Storage and use of data	Transparency		
	Anthropomorphism	Complex operations			
Recommender System	Personalization	Storage and use of data	Misunderstanding of the use and storage of data	Those valuing this technology are ready to sacrifice privacy to reap its benefits. False/manipulated recommendations can lead to mistrust.	(Schafer et al., 2007); (Knijnenburg et al., 2012); (Pu et al., 2012)
	Facilitates decisions	No prevention of the system using feedback/ rating data	Manipulated feedback/ratings		
		Continuous analysis of user behavior	bias and compromised identity		
Predictive Analytics/	Identifies patterns	Storage and use of data	Users are unaware that this ML artifact is used in their technology. Due to this, there	Due to unawareness, there is no reported feedback on the impacts on trust. This	(McCue, 2006); (Glenn & Monteith,
		No prevention in the system			

ML Artifact	Features	Privacy and Security Challenges	Barriers of use	Impacts on Trust	Reference
Big data analytics		using feedback/rating data Continuous analysis of user behavior	is no reported feedback on Barriers to use.	coincides with the ethical concerns of AI.	2014); (Mai, 2016)

A chatbot application is a machine learning application known to simulate discussions with humans, which improves over time by constantly learning algorithms from existing data (Moriuchi et al., 2021). For example, daily use of chatbots might include requesting a support ticket for a refund or exchanging an item purchased online. The chatbot with pre-defined conversations can direct the customer to the correct department or even already process the refund or exchange. Recent research also has shown that chatbots are useful tools to support the online purchase decision making, assisting customers during the (pre)purchase and after-sales support and, particularly, perceived usefulness and rewarding interactions are associated with customer engagement with chatbots which positively affects repurchase intentions (Torres & Delgado, 2022).

Like chatbots, virtual assistants or smart assistants are typically composed of three main components, sensors monitoring environmental and physical characteristics, controllers interpreting the information they receive, and actuators carrying out actions (Lau et al., 2018). Known virtual assistants were developed by companies such as Google, Apple, and Amazon, with capabilities like telling you the weather forecast, finding information about a restaurant, helping you plan your agenda, playing music and podcasts, entertaining, and even telling jokes.

Recommender systems are different machine learning applications using software tools that provide suggestions for items a user may find helpful. Recommender systems are primarily intended for those lacking the personal experience or expertise to assess various alternatives (Soediono, 1989). Recommender systems are used, for example, in e-commerce websites that use individuals' previously tracked purchases to predict and suggest future purchases.

Another Machine Learning application is predictive analytics which forecasts future outcomes by combining historical data with statistical modelling, data mining techniques, and machine learning. Businesses use predictive analytics to spot dangers, opportunities, and future trends by analysing patterns in the data (*What Is Predictive Analytics?* | IBM, n.d.). Daily use of predictive analytic tools might include users accessing navigation maps to discover the best routes and estimate arrival times.

2.2 Trustworthy AI

AI can potentially bring significant benefits to both users and companies that use this technology. However, we are also confronted with challenges with the continuous use of AI applications. Examples include the impact on labor and the use of untrustworthy and unreliable applications that gather individuals' data and information. To build, deploy, and operate AI systems without fear, trustworthy AI focuses on what principles should be implemented (Wickramasinghe et al., 2020). In AI-enabled customer service, the trust concept is even more complex because trust goes beyond technology and brand, including the process and aim of using AI (Hengstler et al., 2016; Siau & Wang, 2018). Research has shown the complex nature

of AI use (Dwivedi et al., 2021) and the uncertainty in several areas, such as security, reliability, and ethics.

Although AI satisfies every criterion of the logical account of trust, it was found that this is a type of reliance rather than trust, as it does not possess emotional states or can be held responsible for its actions (Ryan, 2020). Basing AI as being trustworthy is not compelling because doing so distorts the nature of human relationships, humanizes artificial intelligence, and absolves those who develop and use it of responsibility (Ryan, 2020). However, Glikson & Woolley (2020) showed in their study that (1) the type of trust people build is significantly influenced by machine intelligence and differentiates types of trust depending if it concerns robotic AI or virtual and embedded AI; (2) the degree of machine learning may influence how trust grows, with use and experience showing that a high level of intelligence fosters more trust and, (3) high machine intelligence offers the chance to either preserve the initial high levels of trust (by living up to the high expectations) or minimize the decline of trust for embedded and virtual AI. Also, those authors stated that the trust trajectory for robotic AI is similar to trust development in human relationships, starting low and increasing after the direct hands-on experience. On the other hand, it was observed that significant initial trust declines because of experience with virtual and embedded AI (Glikson & Woolley, 2020).

From a more end-consumer perspective, Ameen et al. (2021) analyze that integrating AI in shopping can lead to a better AI-enabled customer experience. Their findings highlight the central role the concept of trust plays in AI-enabled customer experience, identifying major sacrifices consumers may face, such as a lack of human interaction, loss of privacy, loss of control, time consumption, and possible negative feelings of irritation, all of which can damage AI-enabled service experiences. The authors conclude that when a company offers high-quality, personalized, and convenient AI-enabled services, customers start to trust the brand and its technology. A high level of trust positively impacts service quality.

Hengstler et al. (2016) state that AI technology must be trusted for applied AI to work, the company's role is essential, and communication must also be trusted in the process. As a result, businesses need to start developing trust as technology develops. Significant challenges that emerge from the trustworthiness of AI are personal or organizational privacy and security concerns, the use of biased systems, and even the role AI can take in decision-making processes. A critical AI challenge is guaranteeing that data has consistently high quality since the sources and volume (web, social media, mobile devices, sensors, and IoT) have increased considerably (Anica-Popa et al., 2021). Importantly, in the case of ML algorithms, the underrepresentation of the population in the data used to train the model causes a built-in bias, a vulnerability usually imputed to AI (Shneiderman, 2016). Therefore, trustworthy AI also concerns ML algorithms that tend to incorporate the biases of their human developers, which can induce discrimination in the algorithms and, thus, racist or sexist chatbots (Anica-Popa et al., 2021). For example, most digital assistants are pictured as young Caucasian women, reflecting the stereotyped perception of women in administrative roles (Spencer et al., 2018).

Also, ML models may be seen as “black boxes” because of algorithms' opacity, thus making it difficult to explain how they arrive at certain predictions or suggestions (Adadi & Berrada, 2018; Miller, 2019). One example is using past consumer behaviours as a base for sales forecasting, which may carry on a bias because it is no longer relevant to a customer's current buying need or preference (Anica-Popa et al., 2021). Thus, the development and training of AI algorithms must be responsible and accurate to avoid costly mistakes.

Moreover, ethical principles need to be implemented when deploying algorithms to foster pricing and promotions since retailers could use AI to socially discriminate against customers by using different messages and prices based on customers' specific profiles (Gerlick & Liozu, 2020). From this perspective, AI emerges as an effective marketing tool and a powerful manipulation tool. However, there is a considerable reputational risk that AI-enabled aggressive strategies can generate that could provoke major financial and operational problems (Anica-Popa et al., 2021).

In discussing AI trustworthiness and impacts on individuals and society, critical areas such as medicine and healthcare are especially important since AI solutions or applications can have negative impacts. The European Commission has acknowledged this topic and has proposed ethical guidelines for trustworthy AI. According to their study, trustworthy AI should be composed of three main characteristics: lawful, ethical, and robust. Trustworthy AI needs to be lawful, with instructions and guidelines on what should be avoided, done, and permissible. Ethical and Robust AI are closely related. Ethics refers to the compliance to ethical principles and values. Robust AI refers to AI systems where safeguards should be anticipated to prevent unanticipated adverse effects from a technical perspective. Depending on their social and environmental context, such systems should operate safely and securely. Each element is needed to build trustworthy AI, as the three principles complement each other in an ideal world and operate close together (European Commission, 2019).

2.3 User Privacy and Security

In both physical and digital worlds, privacy and security are crucial. In the digital world, the notions of privacy and security are becoming more and more salient as users deal with these terms and often consent to privacy policies daily, from purchasing an item online to even downloading a simple mobile application. Consequently, external providers can track and use people data in many ways, including customer transaction data from a CRM system, demographic data gleaned from credit card or loyalty card details, survey data linked to an email address, and even visitor data to stores. Additionally, when one considers social media data, which can be connected to individual-level activity, customer-level data becomes very rich and detailed (Bradlow et al., 2017).

Privacy and Security are interconnected, but they are not interchangeable. Privacy can be defined as the level of concern a consumer has regarding a potential infringement of their right to prohibit sharing personal information with third parties (Ischen et al., 2020). Tucker (2018) comments that three factors make privacy challenging: (1) Data may exist for much longer than intended due to the low cost of storage; (2) Data can be repurposed again and utilized for different things than what they were initially meant for; and (3) Data for one person may include information about another person without authorization or knowledge. Since one must have confidence that the entities with whom one has entrusted personal information will protect it effectively, privacy is not only a personal concern but also a security issue in general (Heurix et al., 2015).

However, security entails keeping sensitive data and information safe and protected from unauthorized access (Okta, 2022). Accordingly, Kaur et al. (2018) state that “security is the act of safeguarding data and data resources using technology and procedures preparing from inspections, unauthorized access, recordings, etc.” For instance, in China, a local safari park was sued for “violating consumer protection law” after making facial recognition registration mandatory for visitor entrance. The lawsuit focused on the risk of data leaks: “once leaked,

illegal misuse will easily endanger the safety of consumers,” the safari park has since changed its entrance policy to allow visitors to choose between facial recognition or fingerprint recognition. China’s use of facial recognition is incredibly widespread (see coverage in *The State of AI 2018 & 2019*), but there are some signs that privacy concerns are being taken more seriously.

Sensors are widely used in smart cities, hospitals, and homes, which allows for collecting a massive amount of information from informed and uninformed individuals. To address these issues, is essential to have legislation (Zhang & Tao, 2021). To anticipate various security and privacy threats, currently, there is the existence of laws that protect users regarding information privacy and security compliance. Some of the most known in Europe and the United States of America include the General Data Protection Regulation (GDPR): where organizations worldwide are subject to its requirements as long as they target or gather information about EU citizens, the Health Information Portability and Accountability Act (HIPAA): safeguards user’s data held by healthcare professionals, the Fair Credit Reporting Act (FCRA): protects and limits access of credit information and the Family Educational Right and Privacy Act (FERPA): secures student records held by educational institutions and agencies while allowing authorized students and parents access.

Although we benefit from these passed laws, it is important to notice that these laws are, unfortunately, not enough to hinder organizations that continuously push the boundaries and limits of compliance. Heurix et al. (2015) further observe that companies are increasingly breaking the general rule of data minimization, which states that they should only gather and store as much data as is required to efficiently operate their business processes. Some of the largest Internet-based businesses generate money by obtaining, analyzing, and selling as much personal information about their consumers as possible. Thus, it is important to notice that balancing two competing interests is necessary for a data privacy policy. Customers might not use applications related to AI if there is insufficient protection, and excessive regulation could repress innovation (Davenport et al., 2020).

The scope of AI literature is relatively vast, and users’ perceptions of risk and trust with AI cover different research fields. A bibliometric analysis will be performed to map and understand the current trends in AI-enabled customer experience and, particularly, to examine the impact of security and privacy on users' experience with activities that integrate AI technologies. This method broadens the field and gives in-depth insights into users’ perceptions of risk and trust concerning AI applications. Following, we describe the method analysis and data collection.

2.4 Bibliometric Analysis of Artificial Intelligence

The main objective of this analysis is to present a map of privacy and security research in terms of intellectual roots (intellectual structure) to contribute to a better understanding of the numerous privacy and security research streams regarding AI and its evolution in recent research. Using a bibliometric analysis approach, this study aims to understand the current trends in AI-enabled customer experience and, particularly, to examine the impact of security and privacy on users' experience with activities that integrate AI technologies. This method enables a broad comprehension of the field and in-depth insights.

2.4.1. Methodology

The literature on the associated dangers caused by adopting artificial intelligence on users' security and privacy perceptions and experiences is analyzed using bibliometric approaches. The bibliometric analysis draws on statistical tools to systematically map and decipher large amounts of scientific data and rigorously assess the contribution and impact of publications in a research field (Donthu et al., 2021).

This study adopts a systematic literature review based on the guidelines outlined by Donthu et al. (2021). The procedure includes four steps: (1) definition of the aims and scope; (2) choosing the technique for bibliometric analysis; (3) data collection (includes data cleaning); and (4) running the analysis and reporting.

The aim is to map the current state of the literature on users' perceptions of risk and trust with AI. Since AI is relatively recent, the scope of the research was limited to articles published after 2010. The technique chosen for the bibliometric analysis was co-word analysis using VOSviewer (Visualization of similarities viewer) to identify the interlinkages between the keywords and, thus, understand the key research areas and linkages of the published articles.

The co-word analysis technique is based on “the idea that the co-occurrence of keywords describes the contents of the documents” (Callon et al, 1991). Also, the interlinkage between keywords is shown by the co-occurrence of keywords in a particular cluster (Van Eck & Waltman, 2010). Vosviewer clusters the keywords occurring together most frequently into groups shown in maps.

To collect the data, we use the Web of Science (WoS) database, a comprehensive and leading scientific database. The search was performed on the title, abstract, and keywords, including articles from 2010 till March 2023. The relevant keywords for analysis were chosen based on the revision of extant research relating AI, trustworthiness (or similar terms), and user experience and perceptions. The WoS search string included the following keywords: “Artificial Intelligence” AND (security OR safety OR confident* OR harm* OR Trust* OR privacy) AND (perception* OR “user experience”). The search retrieved 814 results. The results were then reviewed based on the following inclusion criteria: the language of publication (English); and the type of publication (article, proceeding paper, review article, early access, and book chapters). Following these steps, 752 papers were used for further analysis. Then, the publications were manually checked for relevance, and four articles were removed because they were reports from conferences, workshops, and tutorials.

The screening process resulted in a sample of 748 articles (614 articles and 134 proceeding papers) published since 2010 in 369 different journals and 77 different countries. Also, researchers with different backgrounds have contributed to the field, as the maximum number of papers per author is seven, and more than 2800 authors have contributed. Figure 1 shows that AI research is still evolving and is in its initial stage. The number of publications remained low from 2010 until 2014; however, since 2015 there has been a sudden and constant increase in the number of publications. The huge rise in publications is attributable to the widespread interest among academics in the Artificial Intelligence privacy and security domains. USA and China are the countries with more publications (see Figure 2). As Haenlein & Kaplan (2021) note, technological advances in AI may even cause a new era of global governance as China and the USA emerge as leaders in this field, whereas other international regions, including Europe, are behind. This suggests solid academic research is necessary to support consumer, manager, and public policy decision-making. Figure 3 illustrates the different areas of research

that have addressed this topic. The main research areas that tackle this issue are Computer science, Engineering, Business Economics, and Psychology. Importantly, there is an increasing number of research in the healthcare field.

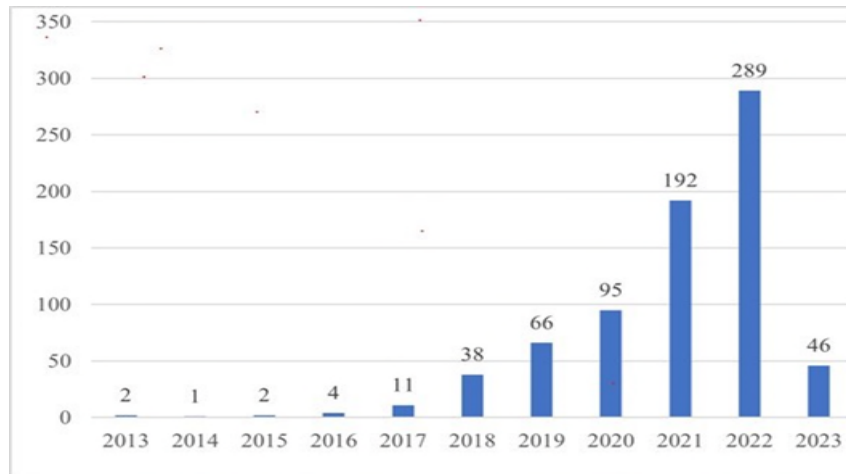


Figure 1. Number of publications published in the period January 2010 - March 2023

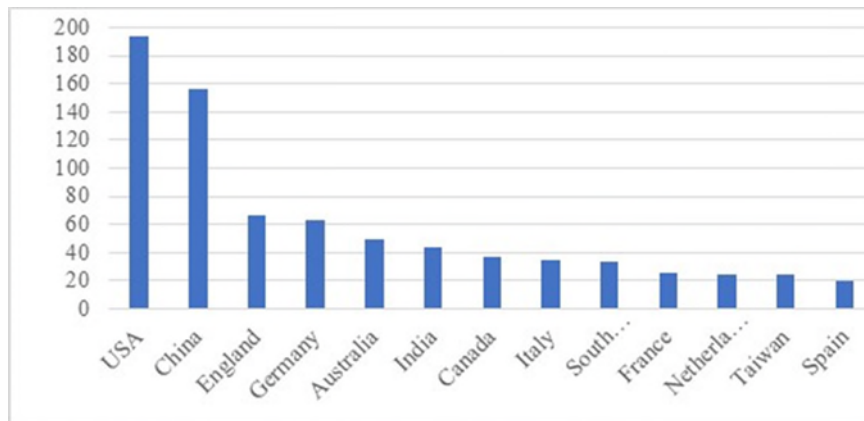


Figure 2. Number of publications per country since 2010 (only countries with more than 20 publications are shown)

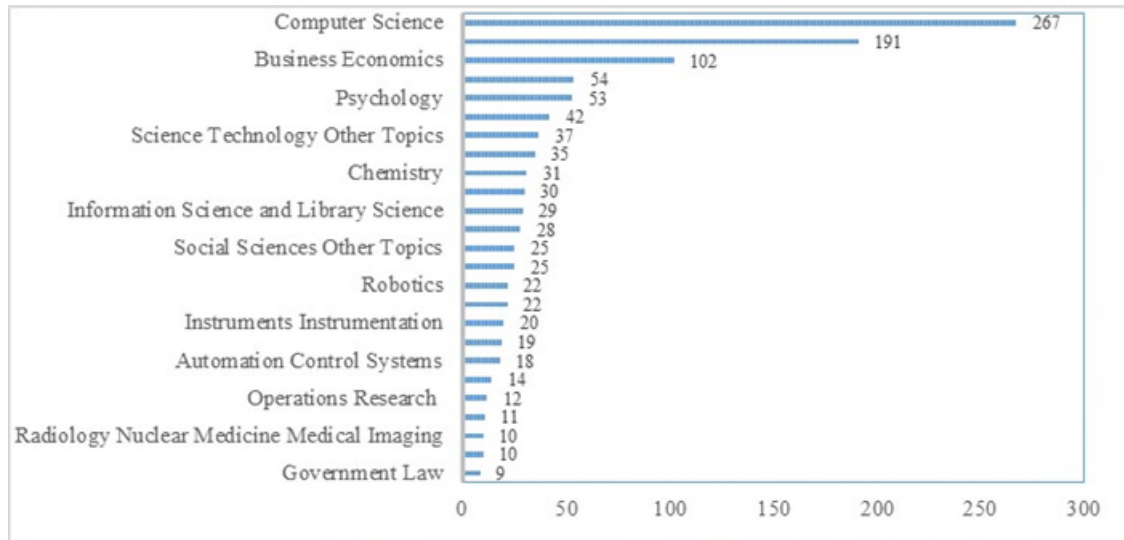


Figure 3. Top research areas by number of publications

To conduct the bibliometric review using co-word analysis, the bibliographic information was inserted in VOS viewer, which used the keywords that appeared most frequently in the title, abstract and author-provided keywords. A thesaurus file was supplied to VOS viewer during analysis to eliminate keyword inconsistency (e.g., singular vs. plural keywords, UK vs. USA English spelling). The network was generated by including keywords that occur at least four times in the sample of articles, which resulted in filtering 3660 keywords to 243 that met the threshold. The keywords irrelevant to our study were deleted, resulting in a final of 241 keywords.

2.4.2 Results Analysis

The papers address many diversified topics; as such, the network created by VOS viewer helped us to understand the emerging research themes and relationships among keywords related to AI and accelerated concerns about trust and safety. Based on the analysis, eight clusters were formed. Table 2 shows the top keywords with more than 30 occurrences. The ones with the highest occurrence are artificial Intelligence, trustworthiness and trust, perceptions, machine learning, acceptance, automated vehicles, robots, and adoption.

Table 2. Top keywords with the highest frequency

Keywords	Cluster	Occurrence	Cluster colors
AI	5	394	purple
trustworthiness and trust	4	159	yellow
perception	1	113	red
technology	7	68	orange
system	1	62	red
machine learning	1	57	red
acceptance	2	55	green
automated vehicles	1	47	red
robot	3	47	blue
adoption	2	43	green
privacy	6	41	grey

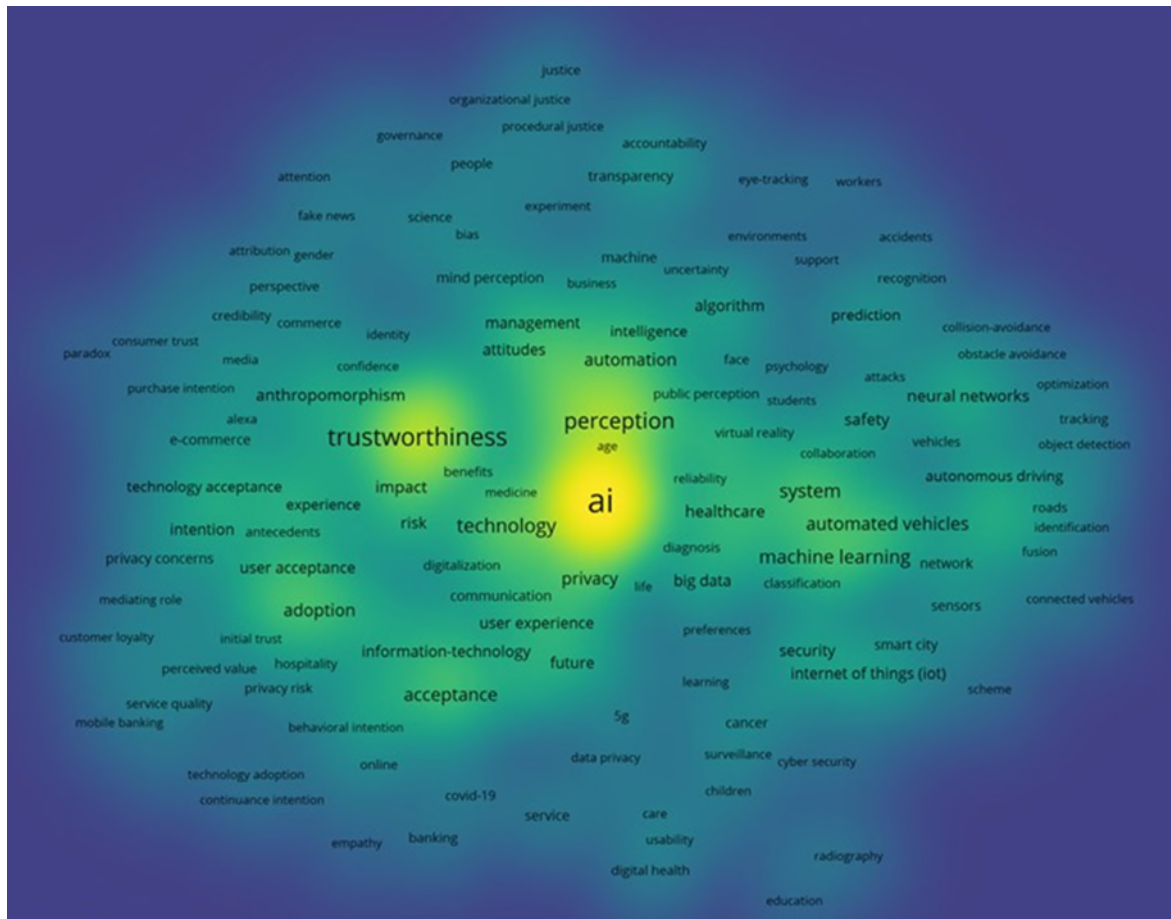


Figure 5. Density visualization of keywords

Table 3. Clusters' themes and sub-themes

Cluster	Theme	Sub-themes
1	AI systems and perceptions	Machine learning, deep learning, and neural networks Automated vehicles and autonomous driving Healthcare Systems perceptions, safety, psychology Biometric data, face recognition, eye-tracking Collaboration, human-robot collaboration
2	AI acceptance and adoption	Risk perception and privacy concerns Perceived value, satisfaction, and loyalty Adoption, technology acceptance, ease of use E-commerce, chatbots, conversational agents
3	Information, experience, and human-machine interaction	Robots, agency, behavior, experience, interactions with humans, dehumanization Challenges, consequences, mind perception, machines Explainable AI
4	Trustworthiness	Automation, performance, measurement Management, accountability, algorithmic decision-making

		Attitudes, transparency, justice, fairness, uncertainty
5	AI	Decision-making, errors Cyber security and ethics AI applications in healthcare Services
6	Internet of Things	Security and privacy Big data Blockchain Cloud computing, edge computing, architecture
7	Technology and anthropomorphism	Anthropomorphism, social responses, personalization, confidence, user perception Willingness, antecedents, online Technology, service robots, automation bias, benefits
8	Expectations and preferences	Expectations, preferences, and personality Competence, universal dimensions

In conclusion, privacy and security are increasingly becoming more critical as technology advances, and academics are now focusing on this issue as the digitalization of services continues to increase. Nonetheless, despite the several benefits for businesses and customers, AI entails several privacy and security risks and challenges for companies and political policies.

The ethical concerns of trustworthy AI are critical to any business system using AI-driven customer profiling. The amount of data collected, tracked, and used by companies enables them to profile customers exploiting their preferences, strongly influencing their buying decision process, and persuading them to purchase. From a customer perspective, these practices entail several risks and increase fears concerning trustful AI applications, affecting companies' brand image. Therefore, from a business perspective, such aggressive strategies produce considerable reputational risks to brand trust that could damage the benefits associated with AI technologies. The key organizational drivers of AI-enabled customer experience investigated by this bibliometric analysis - trust, security, privacy, transparency, and ethics – could be used to manage AI perceived risk and trustworthiness.

Finally, as AI adoption grows, regulators give developers more to think about. Many actors attempt to define principles for the responsible use of AI: common themes include transparency, auditability, robustness, safety, and fairness.

In sum, organizations need to address the following challenges: (i) Ensuring an appropriate level of explainability of AI applications; (ii) Using data of good quality, and (iii) Conducting rigorous model validation.

3 Methodology

In order to address the research questions, this section outlines the research methods employed in this study. The section will be divided into three parts, each corresponding to different phases of the research methodology. A two stage qualitative research approach was chosen, focusing on feedback responses obtained from an online platform as well as conducting semi-structured interviews. In a first stage, it was chosen to focus on the responses obtained from the online platform to obtain impartial insights of the perceptions of users of ML artifacts. However, due to insufficient qualitative data, and in a second stage, semi-structured interviews were carried out to obtain detailed insights of users' perceptions of the use of ML artifacts.

3.1 Data Source and Criteria for Selection

To address the first research question: "What are user's trustiness perceptions of Machine Learning Artifacts?", Google Play Store was utilized as the data source for retrieving user feedback responses. In an attempt to gather as much feedback regarding machine learning applications, it was searched for wide known applications that most users would know, with a minimum of ten million user downloads, more than four-star rating (on a scale from 1-5) and with at least more than a hundred thousand reviews, that fit into the machine learning artifacts categories and features found in Table 1. From this search, four relevant applications emerged and are presented in Table 4:

Table 4. Data Source of ML Artifacts

Machine Learning Artifact	Application Name	Version and n° Downloads (as of 5 th May 2023)	Star rating and n° of reviews	About the Application Quotes
Voice Based Digital Assistant/ Personal Assistant/ Smart assistant	Google Assistant	Version: 0.1.474378801 Downloads: 1B+	Star rating: 4.0 N° reviews: 817K	"Google Assistant is an easy way to use your phone and apps, hands-free"
Chatbot	Bing: Chat with AI & GPT-4	Version: 25.0.41042530 Downloads: 10M+	Star rating: 4.3 N° reviews: 353K	"Experience the safest, most powerful way to search, chat, and find inspiration in a single app. Built on the power of ChatGPT & GPT-4."
Recommender System	Netflix	Version: 8.67.1 build 11 50414 Downloads: 1B+	Star rating: 4.3 N° reviews: 14.1M	"The more you watch, the better Netflix gets at recommending TV shows and movies you'll love."
Predictive Analytics/ Big data analytics	Google Maps	Version: 11.77.0301 Downloads: 10B+	Star rating: 4.1 N° reviews: 17M	"Get real-time GPS navigation, traffic, and transit info, and explore local neighbourhoods by knowing where to eat, drink and go - no matter what part of the world you're in."

To have a deeper understanding of RQ1 in order to better understand the AI trustiness of user perceptions and to respond to RQ2: "What are the benefits and the challenges organizations

face, when implementing these technologies?” And RQ3: “What strategies enhance user trust, address concerns and ensure responsible and secure utilization of these technologies?”, the author conducted in-depth semi-structured interviews to ten general users and ten experts in the field of AI. These interviews occurred between the 14th of May, 2023 and the 6th of June, 2023. As for the general users, the only criteria used for selection was to be 18 years or older. As for the experts, the same criteria apply from the users, and additionally, each expert had to have a background or work in areas of Information Technology (IT). Due to this standard, this particular group divides into 5 interviews done to university professors and 5 interviews done to experts that work in the IT area. A table with demographic data of each group is presented in Table 5 and Table 6.

Table 5. Demographic data of Users'

	Age	Profession
User 01	46	Logistics Analyst
User 02	52	Bus Operator
User 03	18	Student
User 04	20	Student
User 05	38	Restaurant Manager
User 06	29	Personal Trainer
User 07	22	Serving Staff
User 08	19	Kitchen Assistant
User 09	47	Accountant
User 10	18	Student

Table 6. Demographic data of Experts

	Age	Profession
Expert 01	23	Software Engineer
Expert 02	30	System Engineer
Expert 03	28	Software Engineer
Expert 04	58	University Professor and Researcher
Expert 05	48	University Professor and Researcher
Expert 06	23	Software Engineer
Expert 07	44	University Professor and Researcher
Expert 08	37	University Professor and Researcher
Expert 09	23	Data Scientist
Expert 10	58	University Professor and Researcher

3.2 Data Consolidation and Extraction

As each application counted with millions of downloads and user feedback, a web scraping tool known as “Web Scraper” was used to retrieve the most recent data as of 5th May 2023. The web scraping tool enabled to retrieve the date, username and user feedback review from each

Google Play store application listed above in Table 4 and extracted this information into a Microsoft Excel File. All user reviews had to be in English. The Web scraping tool retrieved 2040 reviews from the Google assistant application, 2120 reviews from the Bing: Chat with AI & GPT-4 application, 2120 reviews from the Netflix application and 2000 reviews from the Google maps application. The user feedback reviews span a time period starting from September 12th, 2018, and extend until May 5th, 2023 between all four applications.

The collection of primary data involved conducting semi-structured interviews. All interviewees were contacted by e-mail and requested to participate in an online audio-recorded interview. Upon agreeance, a formal consent form was sent explaining the scope, conditions, confidentiality, and anonymity terms for the study. All data gathered by the researcher from the candidates was anonymous and used informed consent. The interview questions focused on themes such as trust, user experience, benefits and possible concerns associated with Artificial Intelligence technologies. The interviews with experts covered three additional questions, given their expertise in addressing more technical aspects. The interview guide will be included in Appendix A and the consent form provided to candidates will be included in Appendix B.

After the interviews were completed, the voice recordings were transcribed, and certain recordings were translated from Portuguese to English using Microsoft Word.

3.3 Data Analysis

All the data retrieved from the four applications and from the semi-structured interviews were imported into NVivo, a Computer-Assisted Qualitative Data Analysis Software, for analysis purposes and treated in a separate analysis. In the case of Google Play Store reviews, each Excel tab associated with an application was imported as a survey and examined for the content of each column. The usernames and dates were classified as closed-ended questions, while the reviews themselves were categorized as open-ended questions. This categorization facilitated a thorough distinction between reviews and enabled the coding of each individual review. Following this procedure, the author proceeded to perform a text word search query aimed at comprehending the concerns that users may possess. The subsequent keywords were entered into the Nvivo system:

priva* OR secur* OR Transparen* OR Cautio* OR Scar* OR Affraid OR Data OR Concern* OR Listen* OR Watch* OR risk* OR Expos* OR Collect* OR fear OR inform* OR Hack* OR Breach* OR leak* OR voice OR accura* OR danger* OR Safe* OR Harm* OR Worr* OR Wary OR record* OR Vulnerable OR Trust* OR Violate* OR Share* OR Susceptible OR Protect* OR prudent OR Intrusive OR Hear* OR Exploit* OR Careful OR bias*

The results from this text word query resulted in the coding of 1044 reviews from the Google assistant application, 336 reviews from the Bing: Chat with AI & GPT-4 application, 926 reviews from the Netflix application and 542 reviews from the Google maps application.

Then the reviews were manually coded. The ones that did not fit into the context of the search (e.g., reviews referring to the functionality of the app) were removed. The final sample resulted in 94 reviews from the Google assistant application, 112 reviews from the Bing: Chat with AI & GPT-4 application, 45 reviews from the Netflix application and 128 reviews from the Google maps application. The reason for such a significant reduction of reviews for the Google Assistant and Netflix applications, is due to the presence of specific words in the reviews that were also used in the text search query, namely "Voice" for Google Assistant and "Watch" for

Netflix. Because these words are associated to primary features surrounding these applications, users often referred to these words in their reviews. Thus, as most of these reviews did not fit into the context of this study, there was a substantial decrease in the total number of reviews for both applications.

Additionally, a word frequency query was conducted on all four final applications reviews, to understand the current themes that were discussed by the users. The word frequency query was limited to the 1000 most frequent words and each word had to have a minimum length of 5 letters. The query was also grouped by exact matches, words with stem and synonyms.

The primary data from the semi-structured interviews was coded using thematic analysis. Thematic analysis is a qualitative method research approach that can identify and analyse data to report and present main themes that were found in the data that was provided. Thematic analysis is divided into six phases, the first phase being that the researcher familiarizes himself with the data, the second is to generate initial codes, the third is to search for familiar themes in the data, the fourth is to review the themes created, the fifth is to define and name those themes and the sixth is to produce the report with compelling and final extracts that relate to the researchers research questions (Braun & Clarke, 2006).

Based on extant research the following themes were used: benefits, concerns or risks, disadvantages, sensitive data and machine learning artifacts. The coding gathered are further described in Appendix C.

Additionally, a cross tab query was conducted which shows coding intersections between codes and attributes or cases. It was found to be useful to have a distinction between the two groups: general users and experts, so each interviewee was attributed to a “level of expertise” which aided in separating the groups and drawing conclusion that will be discussed below. The cross tab queries, even though extremely useful, did not provide interesting graphs to visualize data, therefore, Microsoft Excel was used to further analyze the cross tab queries from Nvivo.

4 Results

In this section, the results are presented and discussed. First the results from the secondary data are presented followed by the in-depth semi-structured interviews results.

4.1 Results from Google play store reviews

Following the Data Processing and as mentioned, with the help of the text word search query, the author was able to extract 94 reviews from the Google assistant application, 112 reviews from the Bing: Chat with AI & GPT-4 application, 45 reviews from the Netflix application and 128 reviews from the Google maps application that focused on concerns of users. The reviews were analysed using a Word frequency query to identify recurring themes among users. The findings from this analysis are presented in Table 7. Due to the projects limited timeframe, and query's high number of frequent words (more than 1000), only the 10 most frequent occurring words from the Word frequency analysis were chosen. A Word frequency cloud for each application is presented in Appendix D. Furthermore, all concerns coded to each word and classified into potential applicable concerns were examined. To provide context for each applicable concern, real excerpts from testimonials addressing those concerns are presented.

Table 7. Concerns from Customer Reviews

Application Name	10 most frequent words: descending order	Customer Reviews: Excerpts of Testimonials Addressing Concerns	Applicable Concerns
Google Assistant	Phone, Voice, Listening, Constantly, Trying, Disable, Information, Search, Activate, Issue	"... Did you know your phone is always recording? If you mention the C word now background apps installed on Google phones record your data, location and screen and send it off to be processed by third parties, prob the gov!!?"	Fear of being watched
		"Suddenly makes unwanted suggestions by picking a word or two from my private conversations, even when the phone is not in use and just sitting on the table."	Fear of being listened
		"It always wants me to share more information than I'm comfortable with."	Concern of use or collection of Data
Bing	Search, Answer, Microsoft, Information, Biased, Question, using, Privacy, Points, Rewards	"Unfortunately, the language model is very restricted and biased."	Concern for Misinformation or Biased information
		"Thought I'd try the new AI function. Unfortunately, it's too slow to use and often bumps into a brick wall. It also provided inaccurate information"	Concern for Misinformation or Biased information
		"... And no, I am not going to create an account on your server to track me just so I can keep going around and around with your chatbot."	Fear of being watched
		"Not looking forward to an AI bot constantly active for voice."	Fear of being listened
		"... What it did do is immediately transmit 14 pieces of personal information from my phone to a company called Adjust. Uninstalled immediately."	Violation of personal information
Netflix	Change, Watch, History, Documentaries, Information, Always,	"... stop defaulting me to "Everyone's Watching" when it's full of garbage that you have enough data to know I won't watch."	Concern of use or collection of Data

Application Name	10 most frequent words: descending order	Customer Reviews: Excerpts of Testimonials Addressing Concerns	Applicable Concerns
	Quality, Allows, Movies, Security	“Their documentaries contain false information and demonize a certain political party and race.”	Concern for Misinformation or Biased information;
		“Netflix is now invading my privacy and illegally tracking me and anyone else who downloads this app.”	Fear of being watched
		“There have always been issues with the recommendation feature. If you don’t go in and specifically like a video it won’t adjust your preferences, no matter how many times you watch a show. The new set up is horribly limiting and only allows you to see a few movies in each genre, and if you DONT have your preferences perfect you’re never going to see ANYTHING you like.”	Concern for Transparency; Concern of Accuracy
Google Maps	Directions, Driving, Location, Dangerous, Accurate, Change, Options, Minutes, Times, Roads	“Estimated Time of Arrivals that are way off, areas registered as "blue" but upon arrival turns to get traffic ... Difficulty to change route even if original route was already deemed inaccessible. Worse, routes that require to violate the law to take”	Concern of Accuracy; Concern for Ethical Responsibility or Accountability; Concerns over safety;
		“Maps has consistently redirected me into high traffic areas because it apparently thinks it knows better than I do. This has occurred so many times I have lost all trust that I will actually get to where I want to be on time and I'm just uninstalling at this point.”	Concern of Accuracy;
		“Google will still allow companies to pay to be placed ahead of my selection. I typed in Taco Bell, and the first location to be shown was Wendy's, because they paid to be placed ahead of what I wanted. Also, when I type in more vague requests like food, and I've selected closest locations, I do not like that companies can pay to bypass my filter. You're already collecting and selling my personal info, so let me see what I asked for.”	Concern over Company actions; Concern for Transparency
		“Google maps Auto asks questions to improve its directions. One of these questions is, "where is the entrance to ____" That's fine and helpful, but it keeps asking that for my daughter's school! In the US, this info shouldn't be available. Let someone get lost on their first day there. Please stop asking me to let you know where the entrance to my daughter's preschool is! This is a safety concern and I can't even count how many times the app has prompted me to answer this question.”	Concerns over safety;

The analysis of the 10 most frequent words from the queries enabled to identify the most recurring words among the four word search queries. Specifically, the word "Search" appears for both Google Assistant and Bing, the word "information" appears for Google Assistant, Bing, and Netflix, and the word "Change" appears for Netflix and Google Maps. Moreover, users express various concerns including fear of being watched, fear of being listened to, concerns over the use or collection of data, biases in the information provided, concerns about accuracy in these AI applications, violation of personal information, transparency concerns, worries about AI autonomy, ethical responsibilities, safety, and concerns about company actions. These findings follow in line with what has been referenced by other authors (Table 1).

4.2 Results from the analysis of interviews

The results from the secondary data from Google Playstore reviews was useful, but additional in-depth data was needed to respond to the study research questions. As such, semi-structured interviews were conducted to understand if there were additional benefits, disadvantages, and concerns over the use of these types of machine learning artifacts. As discussed above in the data source and criteria for selection section, the interviews were divided into two groups: general users and experts. As discussed in the Data Processing section, all interviews were analysed, and references were coded and placed into categories according to what they referenced. These categories are benefits, disadvantages, concerns or risks, sensitive data and machine learning artifacts.

Additionally, it was noticed that the majority of general users inquired were unaware that certain everyday applications incorporate artificial intelligence technology in some form. However, when these users were provided with an explanation of the specific types of applications that employ AI, they demonstrated the ability to identify some machine learning artifacts that they interacted with daily. An additional question enabled to code a category of the most known and used "Machine Learning Artifacts". Results show that users and experts recognized that the most utilized and well known was Predictive Analytics with reference to applications such as Google Maps and Waze. The second most recognized were chatbots, with reference to ChatGPT and Bing AI. The third most popular, were recommender systems with reference to Amazon, Ebay and streaming platforms. And finally, with only four references, virtual assistants were referenced to applications such as Google Assistant, Siri from Apple, and Alexa from Amazon.

Through the analysis of the interview data, six distinct benefits associated with the utilization of artificial intelligence were identified by both users and experts. The findings are depicted in Figure 6, where the responses from interviewees were categorized accordingly. Among the identified benefits, "efficiency" emerged as the most prominent, consistently acknowledged by both users and experts, as one expert expressed: "I think above all, it's a matter of productivity, of more efficient use of time". However, it is noteworthy that the expert group recognized two additional benefits, namely "personalization" and "decision making," which were not identified by the general user group. This indicates a nuanced perspective and deeper understanding of the potential advantages of AI within the expert group.

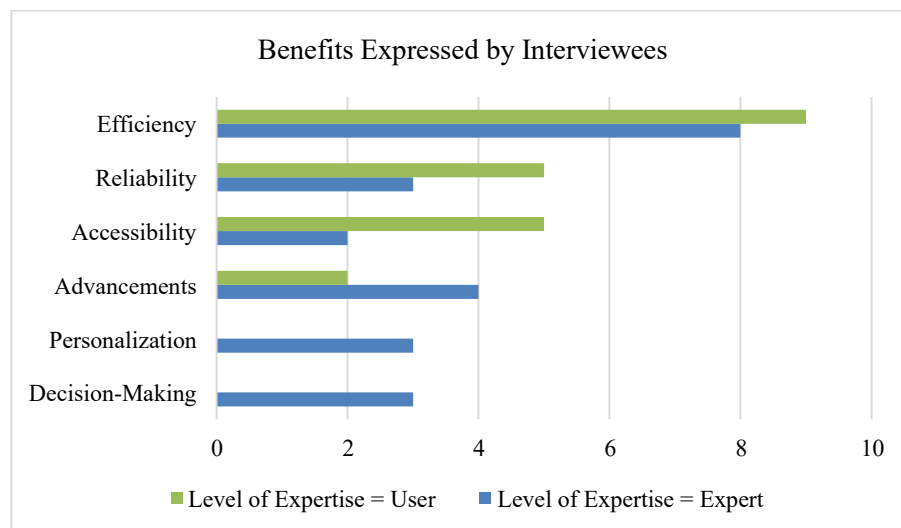


Figure 6. Benefits expressed by interviewees

When considering the disadvantages highlighted by the interviewees, it became evident that several distinct themes were identified. These themes were derived from the interviewee references, and the corresponding codes can be observed in Figure 7. Notably, the expert group identified five additional disadvantages related to the use of artificial intelligence compared to the general users group. However, an interesting distinction arose between the two groups in terms of the emphasis placed on specific drawbacks. Within the expert group, half of the experts pointed out the concern of being misguided or constrained by AI algorithms. On the other hand, the general users predominantly expressed the disadvantage of AI lacking human touch or interaction. These differences are shown in the following quotes:

“There is clearly this problem, because people follow what is recommended, and then end up not questioning what is happening. And we're again being catalogued in a type of client, and this is a problem.” [Expert, Male, 37 years]

“I don't believe that robots can detect the human touch at the end of the day. So you know, as much as you can programme the robot to do things, but there's certain things that you do need human intervention.” [User, Female, 44 years]

“Everything's becoming very systemized, very robotic.” [User, Female, 29 years]

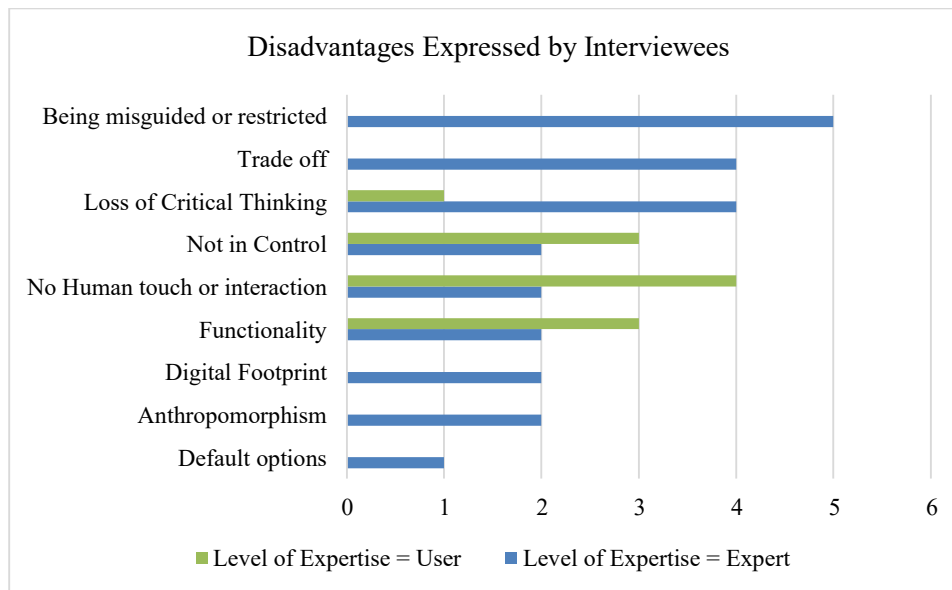


Figure 7. Disadvantages expressed by interviewees

Results showed a distinction between the two groups regarding the factors that made them personally trust these AI applications. Even though most general users mentioned that they did not trust the applications 100%, they tend to ignore this fact and mentioned the benefits associated to its use and acknowledged that the widespread adoption of these applications by other users also influenced their trust, as pointed out by users.

“Not that I really trust them, but it's a form of information... Your search engine is Google, so whatever you need, it's useful to know that whatever you put in there, you get an answer back.” [User, Female, 44 years]

“I would say, because the amount of people that use it on a daily basis ... so I think it just comes down to how many people use the actual system and the reviews that they get and it's just like the norm of using it.” [User, Male, 19 years]

Experts, on the other hand, mentioned they were critical about the feedback they were receiving from these applications and depending on that feedback, that would increase their general trust in a certain system or application. Experts also mentioned that it was not a question of trust but accuracy, as according to the outputs of these applications, they could understand if the information they were receiving was accurate. Also, they noted that due to their knowledge of the limitations and capabilities of these types of AI, they were able to identify more easily situations where they needed to be suspicious of results, as shown in the following quotes.

“I think that trust, you gain with the way the results are being given, if what the answers of the systems whatever they are, in fact we evaluate as being correct and useful.” [Male, Expert, 44 years]

“I know the limitations and capabilities that these tools have. ... But I trust, because I know what's there and when I do, I know why I'm suspicious.” [Expert, Female, 58 years]

Regarding concerns with the use of these AI applications, results are shown in Figure 8. Some concerns mentioned by the user feedback from reviews presented in Table 7 are also mentioned by the interviewees, although additional concerns emerged from the interviews. These additional concerns are related with user unawareness, monopolistic power and the robustness of the systems were solely identified in the expert group. Although, additional concerns such as the concern for the future, data ownership, not exposing personal information and risk of an information breach/leak/hack were identified by both interviewee groups, they were not mentioned in the user reviews from Google Playstore. The quotes below exemplify those concerns.

“In this type of artificial intelligence, I am thinking of technical problems, which cause the systems to not be robust”. [Expert, Male, 48 years]

“I think it's also quite scary is the way you can easily get hacked these days into your bank accounts or things like that. So that is a very scary part.” [User, Female, 44 years]

“At least from my point of view, I'll try to not expose too much.” [Expert, Male, 30 years]

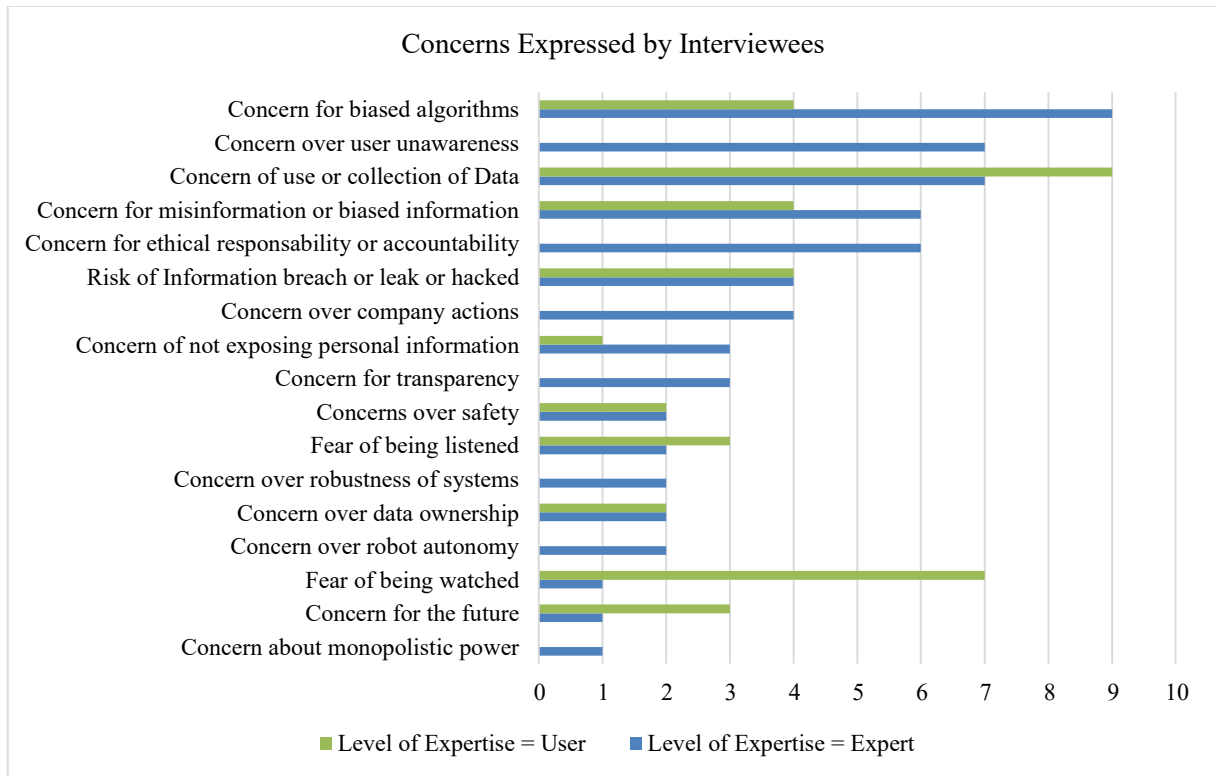


Figure 8. Concerns expressed by interviewees

Results also show that users are not willing to share sensitive information with AI applications or systems. Some interviewees mentioned that bank details, family member information, location, personal addresses, phone numbers and photos are sensitive information that they are reluctant to provide. However, some of the experts also expressed that sharing certain types of information such as medical information and information regarding their personal assets would be off limits. A male expert stressed that “if you ask me information from my medical record, then I find it to be suspicious, isn't it?”

Most of the interviewees mentioned that they are comfortable with AI systems collecting and analysing personal data to improve the system or applications performance, with only 20% saying that they were unsure or stated that it depended on the situation. Whereas for the expert group, 40% would be willing to share their personal information, and half of experts responded that it depended on the situation. Most experts agreed that if the use and collection of this data was used for the overall benefit of society or population such as in terms of advances in the medical field, for example, or other critical areas, it would be considered acceptable as pointed out by one male expert, “It depends a lot on the areas of application. If it's for the common good I have no problem with that”. Although, one male expert had a strong position against sharing personal information saying that “If they are using my data, it is for their own benefit.”

Results show that not all interviewees were willing to share personal information with AI systems to obtain a more personalized experience. It seems that these perceptions are dependent on users’ self-personal perspective and preferences. Interestingly, most users and experts stated that they did not want to feel restricted to specific content and would like to see content out of their normal preferences, as shown below.

“... it doesn't let you explore because it's just constantly feeding you with things. The things you want to see instead of what you might want to actually see, or what you should be looking for.” [User, Male, 22 years]

Both experts and users expressed mixed feelings regarding trust an AI system developed by a government agency or a well-known company, some were unsure, while others bounced between trusting more a well-known company, government agency, both, or none. Again, the results show that personal characteristics and preferences play a critical role in AI trustiness.

Experts identified further concerns and challenges that general users are not aware of. Most experts place an emphasis on legislation and regulation as one of the main challenges for consumers. One expert said that “I would say it's very important for companies to respect the rules in a way. The problem is that sometimes these rules are not well defined.” This is because without proper legislation and regulation, users may be subject to unethical, abuse and non-standardized algorithms that may cause harm to the user. Importantly, legislation and regulation need to keep up with technological advancements to continuously protect consumers. One expert said that “Although regulation is always in the wake of technological innovations, it cannot keep up with the pace of technological executions. So that's one aspect.”

User unawareness is also a big concern, one female expert stated: “People are not educated, they are not trained to handle technology well.” Due to users not understanding the limitations and risks of these types of AI, this might, again, cause user harm, by increasing user bias, the lack of consent and informed decision-making, and diminish overall trust and confidence. A critical concern are the biased algorithms. Experts underlined that, biased algorithms can also cause discrimination and reinforce stereotypes, leading to exclusion and inequality in society, as shown in the quote below.

“There is a version of ChatGPT ... that creates a new context and creates a drawing. If you ask, for example, flight attendant, the drawing that appeared was all women. If you said, manager, without specifying the gender, a man would appear.” [Expert, Male, 58 years old]

The concern for systems robustness and their evolution was also underlined by experts, describing that systems need to be in constant evolution to uphold future advances and developments that are also needed for the future and society. For example, one male expert said that “ChatGPT seemed to have some controversial responses, but over time it was evolving and was becoming more robust.”

Other concerns were mentioned by fewer experts, such as need for reporting abuses that can only happen if companies provide measures to report violations, ensuring that the systems being developed are using quality data from users, and need for adequate collaboration between teams to ensure that the needs of users are being respected and that system are developed correctly by following official guidelines.

5 Discussion

In this section, the author will discuss and draw conclusions from the results obtained from the qualitative study. The objective of this chapter is to respond to understand the users' trustiness perceptions, understand the benefits and challenges that organizations face and ultimately draw strategies that they can follow.

Table 8, represents all the concerns expressed by all sources of information: reviewers, users and experts. Some concerns raised by the google play store feedback reviews, identified in Table 7, were integrated with the users and experts' concerns. Two concerns from table 7, namely "Violation of personal information" and "Concern for Accuracy" have been integrated into the "Risk of Information breach or leak or hacked" category and the "Concern over robustness of systems" respectively. Table 8 intends to depict concerns into various categories for a general understanding as well as describe and understand the concerns stated by reviewers and interviewees to find commonalities.

The results demonstrate that all feedback obtained from the three sources, reviewers, users and experts, commonly agree that privacy and security in general, are the main concerns when using Machine Learning artifacts. Even though users in general do not understand what AI applications entail and experts express a concern over user unawareness in terms of digital literacy, there is still a common perception that these AI applications cannot be fully trusted in terms of their personal privacy and security as well as of the data they provide these systems. From this analysis we can see that 6 concerns out of the 17 have been expressed by all sources and the category "Personal Safety and Privacy" being the only category referenced by all sources. This comes as no surprise. In comparison to previous studies (Table 1), our results demonstrate that there are common concerns supporting that these are real challenges in terms of privacy and security. Another common theme between the three groups is the concern over misinformation or biased information. This shows that users are confronted with these types of information, instilling a sense of distrust. From the results, the three groups demonstrate that they are generally aware and that this is a concern when using these AI applications.

Table 8. Overview of the concerns from the three sources

Category	Concern	Reviewers	User	Expert
Data Privacy and Security	Concern of use or collection of Data	✓	✓	✓
	Risk of Information breach or leak or hacked	✓	✓	✓
	Concern of not exposing personal information	✗	✓	✓
	Concern over data ownership	✗	✓	✓
Information Trust and Integrity	Concern for biased algorithms	✗	✓	✓
	Concern for Misinformation or Biased information	✓	✓	✓
Personal Safety and Privacy	Fear of being watched	✓	✓	✓
	Fear of being listened	✓	✓	✓
	Concerns over safety	✓	✓	✓
Future Implications	Concern for the future	✗	✓	✓

Category	Concern	Reviewers	User	Expert
Corporate actions	Concern over company actions	✓	✗	✓
	Concern for Ethical Responsibility or Accountability	✓	✗	✓
	Concern about monopolistic power	✗	✗	✓
System Reliability, Effectiveness and decision-making	Concern over robustness of systems	✓	✗	✓
	Concern for transparency	✓	✗	✓
	Concern over robot autonomy	✗	✗	✓
Digital Literacy	Concern over User Unawareness	✗	✗	✓

Concerns from the interviewees were synthesised in two tables to better understand users' trustiness perceptions represented by Table 9 and Table 10. To delve deeper into the examination, it was also looked at the number of references for each common concern expressed by the interviewees (Table 9). The column "Total Referenced" represents how many interviewees referenced these concerns in total. And the columns "Referenced by Experts" and "Referenced by Users" specifies how many were referenced by each group. The three main concerns mentioned by both groups are the use and collection of data, the concern for biased algorithms and the concern for misinformation and biased information. However, while users were mostly concerned for the use and collection of data, experts were mostly concerned for biased algorithms. This just enhances further, the difference in expertise and technical perspective between the two groups. Even though experts still express some concern over the use and collection of data, most experts still express high concerns about biased algorithms. The dangers expressed by the experts in the interviews such as: "There is the issue of models, because they are quite complex, they are difficult to understand and for this reason, users are not well aware of the limitations, so they simply accept or can accept all the biases that the model has behind it.". This shows how users might accept and overlook certain biases which can cause other societal challenges. All other concerns represented in Table 9, except "the fear of being watched", are all closely comparable in terms of referenced between the two groups.

Table 9. Common concerns expressed by interviewees

Category	Concerns	Description	Total Referenced (n=20)	Referenced by Experts (n=10)	Referenced by Users (n=10)
Data Privacy and Security	Concern of use or collection of data	Interviewees were concerned on how their data was collected, used and stored by companies that they provided their information to.	16	7	9
	Risk of Information breach or leak or hacked	Interviewees were concerned over the risk of their personal information being breached, leaked or hacked.	8	4	4
	Concern of not exposing personal information	Interviewees aimed to not expose or disclose personal information.	4	3	1
	Concern over data ownership	Interviewees were concerned about who owned their personal data. Users were generally unaware as to who	4	2	2

		has access to their personal information.			
Information Trust and Integrity	Concern for biased algorithms	Interviewees were concerned for unfair or discriminatory algorithms that might treat them unfairly.	13	9	4
	Concern for Misinformation or Biased information	Interviewees were concerned about receiving inaccurate information, as well as information that favoured one side and did not state all the facts.	10	6	4
Personal Safety and Privacy	Fear of being watched	Interviewees fear that someone, or something may be watching them through electronic devices.	8	1	7
	Fear of being listened	Interviewees fear that someone, or something may be listening to them through electronic devices.	5	2	3
	Concerns over safety	Interviewees were concerned on how these algorithms might have the potential to harm individuals or how they can contribute to unsafe outcomes.	4	2	2
Future Implications	Concern for the future	Interviewees were concerned on how AI algorithms may impact our future in terms of control and advancements	4	1	3

Looking at concerns that result mainly from the expert group represented in Table 10, we can understand that these concerns originate from a more technical perspective. These categories seem to express the technical elements of AI systems as well as how organizations actions may generate concerns. From a service implementation perspective, it demonstrates that organizations have much work ahead to provide consistent and reliable systems that incentive users to use them in a confident manner.

Table 10. Concerns expressed solely by the expert group

Category	Concerns	Description
Corporate actions	Concern over company actions	Experts were concerned for companies practices, such as consumer exploitation, purposely disseminating misleading content to generate attention and prioritizing competition in the market rather than safeguarding consumers in terms of privacy and security.
	Concern for Ethical Responsibility or Accountability	Experts are concerned over the grey area that AI provides ethically. Experts are concerned about accountability to when something goes wrong.
	Concern about monopolistic power	The expert was concerned over how there are a small set of dominant companies and people in this field that can ultimately control the models.

Category	Concerns	Description
System Reliability, Effectiveness and decision-making	Concern over robustness of systems	Experts were concerned about the reliability, resilience and stability of AI systems, that can produce errors, biases and unexpected behaviours
	Concern for transparency	Experts were concerned about how these models lack clear guidelines about consumers personal information and the decision criteria used by the models.
	Concern over robot autonomy	Experts were concerned about AI performing automatic decisions without human validation.
Digital Literacy	Concern over User Unawareness	Experts were concerned about how consumers were not aware of how the systems function and results are obtained from these algorithms.

The analysis of the benefits and disadvantages referenced by the interviewees can be very useful for companies that would like to implement these systems enhancing user experience. The result of this analysis is present in Table 11. From the benefits we can understand that in a service context all of these can positively impact the user experience and improve company performance. The four benefits mentioned by both users and experts, just shows how AI has the capability of transforming our lives in society. As a society, we are always looking for solutions and practices that are accessible, reliable, and efficient, while hopefully, as well, contributing to the future, and AI acknowledges all of these traits. Decision-making enhancements and personalization even though not referenced by the users, can immensely benefit the organization. If used properly, it can result in higher revenues, by creating and responding more efficiently to needs and wants of customers, which can contribute to happier customers as well.

As for the challenges, there is an issue between human and technology interaction. Anthropomorphism, the lack of human touch, the loss of critical thinking, being misguided or restricted from AI and the trade-off of information for potential benefits all raise concerns as to what implications this may have in the future and raises questions on where we can draw a line. Even though advancements in this area are crucial for the benefit of society, it is still unknown what are the impacts of these challenges. For example, the lack of human interaction in a service automation context may affect the user experience, as well as the trade-off of personal information for benefits may raise concerns not only on a personal level but also in societal terms. This is in line with previous research, where governments are using surveillance and facial recognition to ensure “safe” cities.

Also, challenges surrounding the capabilities of the AI systems are in question. Digital footprints raise the challenge for consumers as well as companies, as to where and until when organizations store data and what happens with that data when it is no longer needed. The functionality of systems that raise challenges in terms of accuracy and reliability, the default options that are present in the systems and are hard to understand and unsubscribe, and the users not being able to be in control of what the AI systems present them may all contribute to a negative user experience. Is important to have a clear understanding between the benefits AI can provide and the challenges that organizations may face. With the help of this knowledge, organizations can plan and integrate better strategies, that foresee potential problems.

Table 11. Benefits and Challenges for Organizations based on User and Expert perspectives

		Description	User	Expert
Benefits	Decision-Making	AI tools helps users make easier and better decisions	✗	✓
	Personalization	AI tools have the capacity to provide personalized solutions and experiences	✗	✓
	Future Advancements	AI tools have the capacity to boost innovation and promote well-being in society	✓	✓
	Accessibility	AI tools are easily accessible in terms of learning and understanding and use	✓	✓
	Reliability	AI tools can be relied on to perform tasks	✓	✓
	Efficiency	AI tools can accomplish tasks with minimal wasted time and effort	✓	✓
Challenges	Default options	AI tools may come with default options for users, which are difficult to understand and handle according to their preferences	✗	✓
	Anthropomorphism	Users might attribute human-like features to AI tools that might lead to unrealistic expectations and perceptions	✗	✓
	Digital Footprint	Users are not aware that all their information that is provided online, stays online, leaving a digital footprint	✗	✓
	Functionality	AI tools may not perform according to user expectations	✓	✓
	No Human touch or interaction	AI tools are not capable of interacting as well as a human to human interaction	✓	✓
	Not in Control	AI tools can overpower users and make them feel like they are not in control	✓	✓
	Loss of Critical Thinking	AI tools promote that users rely on their systems, leading to a loss of critical thinking	✗	✓
	Trade off	For users to obtain any benefits from AI tools, there is always a trade-off of information.	✗	✓
	Being misguided or restricted	AI tools might make users feels restricted or misguided by the information they provide	✗	✓
	Ensuring proper legislation and regulation	AI tools might not follow industry standards and practices regarding legislation and regulation of AI	✗	✓

To help create some balance, this study suggests the implementation of strategies, that were based on all the concerns, benefits, and challenges the sources of this study provided. Table 12, demonstrates strategies that companies can ultimately execute when trying to introduce and implement these ML artifacts in their organizations. The compilation of this table stems from deductions and interpretations derived from results, however the table is supported and takes into consideration, the opinions and expertise of experts, showing representative quotes that

can provide more context for each strategy. This table is divided into four parts, the first refers to the context in which the companies can implement their strategies, three contexts were provided: the user, the AI system and in the company introspection or self-awareness. The second and third are the strategy that can be implemented and the description of the strategy. The fourth represents quotes from experts that can further strengthen the strategy position.

Table 12. Strategy implementation for Organizations

Strategy Context	Strategy	Description of the Strategy	Quote from Experts
User	Awareness and Education	Users may not be aware of current dangers and biases of AI. Reinforcing user awareness and education of AI tools that are implemented in organizations is essential.	"I think there are a lot of people who don't quite have the notion of the risks and vulnerabilities. I'm not saying it's important for everyone to know how it works, but it's important for them to know the limitations, to know the risks and to be aware."
	Empowerment	Users should be encouraged to report any abuse online regarding security and privacy violations. Organizations should establish mechanisms for users to report such situations.	"When companies do not respect the code of conduct, consumers should protest and file lawsuits."
AI System	Driving Progress	Continuous system upgrades and evolution are crucial to ensure that systems are robust and secure.	"Some protection in the system would be important without cutting development completely, because there is also this prospect that we should end AI use. And I guess that's not the way to go."
	Fostering Inclusivity	Inclusion measures need to be in place to mitigate bias that are engrained in the algorithms.	"We need to include these kinds of concerns in the early stages of design of the system so that they are less susceptible to the problem of biased data that is used to build it, but this is not easy at all to achieve."
Company Introspection or Self-Awareness	Fostering Synergies	Implementing teamwork and collaboration strategies is important to bring together different levels of expertise and promotes a user-centric approach.	"Developers need to understand a little of the other areas involved in the implementation. They need a bit of multidisciplinary knowledge to be able to address user needs."
	Data Integrity	Ensuring the integrity and excellence of data through rigorous quality assurance measures.	"I think that, above all, whoever develops or trains these models, has to be careful to have quality data"

Strategy Context	Strategy	Description of the Strategy	Quote from Experts
	Upholding compliance	Ensuring the establishment and enforcement of legislative and regulatory measures to promote compliance.	“It is up to employers and companies to make their professionals respect the codes of conduct so that the software and hardware are implemented correctly.”

As for each strategy and starting in the context of implementing strategies for the user, two approaches are considered, the awareness and education of the user and the empowerment of the user. As found in the results and discussed, the users clearly have a lack of awareness of AI systems, which may lead to dangerous situations that can encourage biases for example and promote a loss of critical thinking as users are not aware of the capabilities and limitations of these systems. If organizations implement this approach, it can increase trustiness, as this contributes to users being more aware of how the systems work and what data is being used and collected and can also confront challenges such as anthropomorphism. As for the user empowerment strategy, companies should implement measures to help users report privacy and safety concerns when using their system. This can help address challenges such as the users feeling of not being in control and the challenge revolving default options, as well as addressing their concern for transparency. This strategy can also amplify user feedback, which can contribute to the improvement and evolution of the system and service design.

In the context of applying strategies towards the AI systems, driving progress seems to be an essential strategy in any organization’s strategy implementation, however it is even more important here. Given that these AI tools rely and base their information mostly on user preferences and historical data, it is crucial to keep the systems updated so organizations can maintain the accuracy, reliability, and security of the systems. This addresses concerns of the experts in terms of the robustness of the systems and further highlights the findings from the bibliometric analysis. Another strategy that needs to be implemented in the context of AI systems, should be to foster inclusivity. One of the main concerns for users and experts was about the trust and integrity of information. Organizations need to implement measures that promote inclusivity to fight discrimination and bias. In a service context, by implementing these measures, this can also help build users trust, as it includes them and makes them perceive that their interests and features are acknowledged and are important. By also having a wide range of data it can generate more accurate results.

Organizations also need to be introspective and consider strategies that internally can improve their AI systems. Fostering Synergies, is common practice in organizations, however when developing and integrating these systems in a service contexts, organizations require increased collaboration and teamwork between different teams, due to the fact that these are complex systems that need different perspectives (e.g. marketing, customer service, IT development, etc.). Different perspectives and levels of expertise that take into account different user needs are essential. Another strategy and referenced in the bibliometric analysis, is to ensure data integrity. By ensuring data integrity and confirming data sources, there is less likelihood of false information or biased information circulating. With accurate data, AI systems like these can enhance decision making processes and feel reliable to users. Finally, to ensure ethical and responsible practices, organizations should establish strategies that align with official guidelines and standards that are present in the tech community. This can involve the formation

of, for example, a team responsible for reviewing and adhering to official guidelines. Organizations need to prioritize ethical considerations and privacy and security user concerns over short-term gains, which in the long term can build trust, demonstrate integrity, and foster sustainable growth in the market.

6 Conclusion and Future research

In this section, it will be explored the main conclusions from this research and discussed limitations of the study as well as future research, that can help to provide further insights and advancements in the field.

6.1 Conclusion

AI tools, namely machine learning artifacts such as virtual assistants, chatbots, recommender systems and predictive analytics can be a convenient and supportive tool in any service automation context.

This study makes a significant theoretical contribution. This analysis provided important information of general trends in the field as well as the impact of security and privacy on users' experience. From this analysis emerged the inter-relations of the themes with diverse fields and research areas, as well as the growing research on this topic, allows to conclude that there is a general and increasing concern regarding this topic. It was also noticed, that ethical concerns and the responsible use of AI call for more rules and regulation, as this is of the utmost importance for future research. The key organizational drivers of AI-enabled customer experience investigated by this study - trust, security, privacy, transparency, and ethics – could be used to manage AI perceived risk and trustworthiness (Faria et al., forthcoming).

This study has also successfully addressed the three research questions. The first research question, relating to user's trustiness perception, shows that users have various concerns and negative perceptions around the use of AI tools which is consistent to what has been found by previous authors (Table 1). Also the bibliometric analysis emphasizes the fact of that privacy and security is an increasing concern as technology advances, and that academics are now focusing on this issue, as the digitalization of services continues to increase. It was found seventeen different concerns between reviewers, users and experts and while all sources of information share most concerns, the group of experts were able to identify concerns related to more technical aspects that users are not aware of. However, it was noticed that the category of personal safety and privacy concerns (referenced by all sources), deems to be the most recognized. Delving into a deeper examination from the interviewees, the most referenced concerns in common were the use and collection of data, the concern for biased algorithms and the concern for misinformation or biased information. However, between both groups each one put emphasis on a different category. Nonetheless, despite the several benefits for both businesses and customers, we can conclude that AI entails several privacy and security risks and challenges for companies. This research question also allows to enlighten new concerns such as in the category of corporate actions surrounding concerns over company actions and the monopolistic power of some organizations.

To understand the benefits and the challenges organizations face, when implementing these technologies (RQ2), the study focused on dissecting benefits and disadvantages of the use of AI, that were mentioned solely by the interviewees. By understanding what interviewees most gain and lose when using these AI tools, we can also understand what can become a benefit and challenge for an organization. From the results, six benefits, were able to be identified and ten different challenges were able to be acknowledged. This study also emphasizes the challenge that some companies may face regarding the improvement of their system and the general distrust in companies or government agencies. 50% of the experts would only provide information to improve an AI system if it would benefit society in critical areas, such as

healthcare. This shows that there is a general distrust and that most experts do not believe that most organizations have the best interests at heart. The same goes whether interviewees would trust more a government agency or organization that develops AI tools. There was no general consensus whether they would trust one more over the other which leads us to conclude that there is still a perception of distrust regarding both entities. The main conclusions from this research question, allow organizations to strategically plan and develop AI systems that can gain from the benefits and respond to the challenges in an adequate manner.

Overall companies also need to understand that machine learning artifacts mainly base their results on historical data and learn from experience by identifying patterns in data. Wars, economic or political crises, pandemics and other disruptions that may occur in society can disrupt data availability and quality, for example. Recently the covid-19 pandemic emphasized that user habits, responses and experiences might change rapidly, which may cause some disruption when utilizing these tools.

To address the last research question, this study suggests the implementation of strategies that can help guide organizations that address concerns and challenges. Based on the all three sources of information it was able to identify 7 different strategies that can help organizations when implementing these AI tools. This is one of the main contributions of this study as it serves as theoretical guidance and provides valuable insights from experts that have more knowledge in this area.

6.2 Limitations and Future Research

Future research initiatives can help to address the shortcomings of this study. One limitation of this study is the focus on only four machine learning artifacts. There are other Machine learning artifacts that can also be applied in a service context, such as image and video recognition and machine learning applied to smart home systems and cities. A more comprehensive study can be done to include these other machine learning artifacts.

The fact that the bibliometric data was generated only from the Web of Science is a research constraint. Hence, different data sources, such as Scopus or SpringerLink, should be used in the next investigations. Based on 748 documents, the study provides an overview of various factors, such as published document categories, publication sources, and VOSviewer-referenced research. Furthermore, a more detailed oriented analysis can be performed to understand the impact of each research area found in this current study. This study also focused solely on user perceptions of privacy and security. However, academics can further research the interconnected areas and interrelate them.

Another limitation of this study is based on the source and processing of data. Google play store reviews were mainly focused on the functionality of the application, which constrained our research. In addition, the chatbot application *Bing: Chat with AI & GPT-4*, might not accurately represent chatbots that are used in service contexts. In a service automation context, users interact with chatbots in a different way, and might have different perceptions than those represented in the google play store reviews. To assist this limitation, future research can attempt to locate better samples or research these concerns in an organization context that can provide better user feedback. Regarding the processing of data, it was considered that the text search query might not have found or represented all concerns, a deeper analysis can be made to resolve this issue.

Regarding additional future research, it is suggested that the strategies that were suggested in the last research question, might be implemented in an organization as a case study, to understand its applicability.

7. References

- Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, 6, 52138–52160.
<https://doi.org/10.1109/ACCESS.2018.2870052>
- Ameen, N., Tarhini, A., Reppel, A., & Anand, A. (2021). Customer experiences in the age of artificial intelligence. *Computers in Human Behavior*, 114.
<https://doi.org/10.1016/j.chb.2020.106548>
- Anica-Popa, I., Anica-Popa, L., Rădulescu, C., & Vrîncianu, M. (2021). The Integration of Artificial Intelligence in Retail: Benefits, Challenges and a Dedicated Conceptual Framework. *Amfiteatru Economic*, 23(56), 122–122.
<https://doi.org/10.24818/EA/2021/56/120>
- Artificial Intelligence and Machine Learning in the Insurance Sector*. (n.d.). Retrieved July 3, 2023, from <https://www.strategy-transformation.com/artificial-intelligence-and-machine-learning-in-the-insurance-sector/?cv=1>
- Bellman, R. E. (1978). *An introduction to artificial intelligence: Can computers think?* (Boyd & Fraser Pub., Ed.).
- Benaich, N., & Hogarth, I. (2022). State of AI Report About the authors. *Nathan Benaich Ian Hogarth*, 191.
- Bolton, R. N., McColl-Kennedy, J. R., Cheung, L., Gallan, A., Orsingher, C., Witell, L., & Zaki, M. (2018). Customer experience challenges: bringing together digital, physical and social realms. *Journal of Service Management*, 29(5), 776–808.
<https://doi.org/10.1108/JOSM-04-2018-0113>
- Bradlow, E. T., Gangwar, M., Kopalle, P., & Voleti, S. (2017). The Role of Big Data and Predictive Analytics in Retailing. *Journal of Retailing*, 93(1), 79–95.
<https://doi.org/10.1016/J.JRETAI.2016.12.004>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics*, 22(1), 155–205.
<https://doi.org/10.1007/BF02019280/METRICS>
- Cheng, Y., & Jiang, H. (2020). *How Do AI-driven Chatbots Impact User Experience? Examining Gratifications, Perceived Privacy Risk, Satisfaction, Loyalty, and Continued Use*. <https://doi.org/10.1080/08838151.2020.1834296>
- Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, 60.
<https://doi.org/10.1016/J.IJINFOMGT.2021.102383>
- Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48(1), 24–42. <https://doi.org/10.1007/s11747-019-00696-0>

- de Barcelos Silva, A., Gomes, M. M., da Costa, C. A., da Rosa Righi, R., Barbosa, J. L. V., Pessin, G., De Doncker, G., & Federizzi, G. (2020). Intelligent personal assistants: A systematic literature review. *Expert Systems with Applications*, 147. <https://doi.org/10.1016/J.ESWA.2020.113193>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/J.JBUSRES.2021.04.070>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- European Commission. (2019). Ethics guidelines for trustworthy AI. High-Level Expert Group on Artificial Intelligence. *European Commission*, 1–39.
- Faria, R., Torres, A. & Beirão, G. (accepted, in press). Trustworthy Artificial Intelligence and Machine Learning: Implications on Users' Security and Privacy Perceptions, in "Confronting Security and Privacy Challenges in Digital Marketing, Editors Pires, P., Santos, J., Veiga, I. & Torres, A. Projected Release Date: June, 2023. Pages: 330, DOI: 10.4018/978-1-6684-8958-1. ISBN13:9781668489581|ISBN10:1668489589|EISBN13: 9781668489604
- Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018). What makes users trust a chatbot for customer service? An exploratory interview study. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11193 LNCS, 194–208. https://doi.org/10.1007/978-3-030-01437-7_16
- Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., Moro, E., Wang, D., Youn, H., & Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences of the United States of America*, 116(14), 6531–6539. https://doi.org/10.1073/PNAS.1900949116/SUPPL_FILE/PNAS.1900949116.SAPP.PDF
- Gerlick, J. A., & Liozu, S. M. (2020). Ethical and legal considerations of artificial intelligence and algorithmic decision-making in personalized pricing. *Journal of Revenue and Pricing Management*, 19(2), 85–98. <https://doi.org/10.1057/s41272-019-00225-2>
- Glenn, T., & Monteith, S. (2014). Privacy in the Digital World: Medical and Health Data Outside of HIPAA Protections. *Current Psychiatry Reports*, 16(11). <https://doi.org/10.1007/S11920-014-0494-4>
- Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660. <https://doi.org/10.5465/annals.2018.0057>
- Haenlein, M., & Kaplan, A. (2021). Artificial intelligence and robotics: Shaking up the business world and society at large. *Journal of Business Research*, 124, 405–407. <https://doi.org/10.1016/j.jbusres.2020.10.042>

- Hasal, M., Nowaková, J., Ahmed Saghair, K., Abdulla, H., Snášel, V., & Ogiela, L. (2021). Chatbots: Security, privacy, data protection, and social aspects. *Concurrency and Computation: Practice and Experience*, 33(19). <https://doi.org/10.1002/CPE.6426>
- Hengstler, M., Enkel, E., & Duelli, S. (2016). Applied artificial intelligence and trust-The case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change*, 105, 105–120. <https://doi.org/10.1016/j.techfore.2015.12.014>
- Heurix, J., Zimmermann, P., Neubauer, T., & Fenz, S. (2015). A taxonomy for privacy enhancing technologies. *Computers and Security*, 53, 1–17. <https://doi.org/10.1016/J.COSE.2015.05.002>
- Inman, J. J., & Nikolova, H. (2017). Shopper-Facing Retail Technology: A Retailer Adoption Decision Framework Incorporating Shopper Attitudes and Privacy Concerns. *Journal of Retailing*, 93(1), 7–28. <https://doi.org/10.1016/j.jretai.2016.12.006>
- Ischen, C., Araujo, T., Voorveld, H., van Noort, G., & Smit, E. (2020). Privacy Concerns in Chatbot Interactions. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11970 LNCS, 34–48. https://doi.org/10.1007/978-3-030-39540-7_3
- Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586. <https://doi.org/10.1016/J.BUSHOR.2018.03.007>
- Kaur, P., Sharma, M., & Mittal, M. (2018). Big Data and Machine Learning Based Secure Healthcare Framework. *Procedia Computer Science*, 132, 1049–1059. <https://doi.org/10.1016/J.PROCS.2018.05.020>
- Knijnenburg, B. P., Willemsen, M. C., Gantner, Z., Soncu, H., & Newell, C. (2012). Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, 22(4–5), 441–504. <https://doi.org/10.1007/S11257-011-9118-4>
- Lau, J., Zimmerman, B., & Schaub, F. (2018). *Alexa, Are You Listening? Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers*. <https://doi.org/10.1145/3274371>
- Mai, J. E. (2016). Big data privacy: The datafication of personal information. *Information Society*, 32(3), 192–199. <https://doi.org/10.1080/01972243.2016.1153010>
- Mccarthy, J. (2007). *WHAT IS ARTIFICIAL INTELLIGENCE?*
- McCue, C. (2006). Data mining and predictive analytics in public safety and security. *IT Professional*, 8(4), 12–18. <https://doi.org/10.1109/MITP.2006.84>
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. In *Artificial Intelligence* (Vol. 267, pp. 1–38). Elsevier. <https://doi.org/10.1016/j.artint.2018.07.007>
- Moriuchi, E., Landers, V. M., Colton, D., & Hair, N. (2021). Engagement with chatbots versus augmented reality interactive technology in e-commerce. *Journal of Strategic Marketing*, 29(5), 375–389. <https://doi.org/10.1080/0965254X.2020.1740766>
- Nithya, B., & Ilango, V. (2017). Predictive analytics in health care using machine learning tools and techniques. *Proceedings of the 2017 International Conference on Intelligent*

- Computing and Control Systems, ICICCS 2017, 2018-January*, 492–499.
<https://doi.org/10.1109/ICCONS.2017.8250771>
- Okta. (2022). *Privacy vs. Security: Exploring the Differences & Relationship*. Okta.
<https://www.okta.com/identity-101/privacy-vs-security/>
- Pu, P., Chen, L., & Hu, R. (2012). Evaluating recommender systems from the user's perspective: Survey of the state of the art. *User Modeling and User-Adapted Interaction*, 22(4–5), 317–355. <https://doi.org/10.1007/S11257-011-9115-7>
- Russell, S., & Norvig, P. (2010). *Artificial Intelligence A Modern Approach Third Edition*. In Pearson. <https://doi.org/10.1017/S0269888900007724>
- Ryan, M. (2020). In AI We Trust: Ethics, Artificial Intelligence, and Reliability. *Science and Engineering Ethics*, 26(5), 2749–2767. <https://doi.org/10.1007/s11948-020-00228-y>
- Schafer, J. Ben, Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4321 LNCS, 291–324. https://doi.org/10.1007/978-3-540-72079-9_9
- Shneiderman, B. (2016). The dangers of faulty, biased, or malicious algorithms requires independent oversight. *Proceedings of the National Academy of Sciences of the United States of America*, 113(48), 13538–13540.
<https://doi.org/10.1073/PNAS.1618211113/ASSET/6360B2B9-0BFB-421E-9B5F-A5F7E59B5CF6/ASSETS/GRAPHIC/PNAS.1618211113FIG02.JPEG>
- Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2), 47–53.
- Soediono, B. (1989). Recommender Handbook. In *Journal of Chemical Information and Modeling* (Vol. 53). <https://doi.org/10.1017/CBO9781107415324.004>
- Spencer, D. J., Poggi, J., & Gheerawo, R. (2018). Designing out stereotypes in artificial intelligence: Involving users in the personality design of a digital assistant. *ACM International Conference Proceeding Series*, 130–135.
<https://doi.org/10.1145/3284869.3284897>
- Sunarti, S., Fadzlul Rahman, F., Naufal, M., Risky, M., Febriyanto, K., & Masnina, R. (2021). Artificial intelligence in healthcare: opportunities and risk for future. *Gaceta Sanitaria*, 35, S67–S70. <https://doi.org/10.1016/J.GACETA.2020.12.019>
- Syam, N., & Sharma, A. (2018). Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, 69, 135–146.
<https://doi.org/10.1016/J.INDMARMAN.2017.12.019>
- Torres, A. I., & Delgado, C. J. M. (2022). *Assessing Customer Interactions With Chatbots in Online Shopping Experiences* (José Santos and Bruno Sousa, Ed.; Promoting, pp. 203–223). IGI Global book. <https://doi.org/10.4018/978-1-6684-5523-4>
- Tucker, C. (2018). *Privacy, Algorithms, and Artificial Intelligence*. University of Chicago Press, 197–236.

- van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Vimalkumar, M., Sharma, S. K., Singh, J. B., & Dwivedi, Y. K. (2021). ‘Okay google, what about my privacy?’: User’s privacy perceptions and acceptance of voice based digital assistants. *Computers in Human Behavior*, 120. <https://doi.org/10.1016/J.CHB.2021.106763>
- What is predictive analytics?* | IBM. (n.d.). Retrieved February 3, 2023, from <https://www.ibm.com/topics/predictive-analytics>
- Wickramasinghe, C. S., Marino, D. L., Grandio, J., & Manic, M. (2020). Trustworthy AI development guidelines for human system interaction. *International Conference on Human System Interaction, HSI, 2020-June*, 130–136. <https://doi.org/10.1109/HSI49210.2020.9142644>
- Zadeh, L. A. (2008). Toward human level machine intelligence - Is it achievable? the need for a paradigm shift. *IEEE Computational Intelligence Magazine*, 3(3), 11–22. <https://doi.org/10.1109/MCI.2008.926583>
- Zhang, J., & Tao, D. (2021). Empowering Things with Intelligence: A Survey of the Progress, Challenges, and Opportunities in Artificial Intelligence of Things. In *IEEE Internet of Things Journal* (Vol. 8, Issue 10, pp. 7789–7817). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/JIOT.2020.3039359>

APPENDIX A: Interview Guide

1. What is your age?
2. What is your profession?
3. What is your experience using AI applications?
4. What type of AI applications do you use?
5. What benefits do you identify when using AI Applications?
6. What are the disadvantages do you identify when using AI Applications?
7. What makes you trust these AI applications?
8. Do you have any concerns when interacting with one of those AI applications? Do you have any barriers of use?
9. Ethically, do you have any concerns?
10. How comfortable are you with AI systems collecting and analysing your data to improve their performance? Why or why not?
11. Are you willing to share personal information with AI systems if it results in a more personalized experience? Why or why not?
12. Would you be more likely to trust an AI system that is developed by a government agency or a well-known company? Why or why not?

Additional questions asked to the experts:

13. What do you consider to be the risks and challenges of Artificial Intelligence in the consumer experience?
14. What steps do you think should be taken to ensure that AI systems respect users' privacy and data protection rights?
15. What concerns, if any, do you have about the potential biases in AI algorithms?

APPENDIX B: Consent Form

CONSENTIMENTO INFORMADO, LIVRE E ESCLARECIDO PARA PARTICIPAÇÃO EM PROJETOS DE DOCÊNCIA E/OU INVESTIGAÇÃO

Por favor, leia com atenção a seguinte informação. Se achar que algo está incorreto ou que não está claro, não hesite em solicitar mais informações. Se concorda com a proposta que lhe foi feita, queira assinar este documento.

Título do estudo: People and (Ro)Bots: Trustworthy AI and Machine Learning Applications for Services Automation.

Enquadramento: Dissertação de Mestrado em Engenharia de Serviços e Gestão (MESG) com colaboração entre o INESC TEC e FEUP, com orientação da Professora Doutora Ana Torres e Professora Doutora Gabriela Beirão.

Explicação do estudo: O estudo em curso visa investigar as perceções, atitudes e experiências dos indivíduos em relação à utilização de tecnologias de Inteligência artificial. Será adotada uma abordagem de pesquisa qualitativa, especificamente utilizando entrevistas semiestruturadas. As perguntas da entrevista giraram em torno dos temas de confiança, transparência, responsabilidade, experiência do utilizador, benefícios e possíveis preocupações associadas às tecnologias de Inteligência artificial. A entrevista deve demorar aproximadamente entre quinze a trinta minutos.

Condições e financiamento: Este estudo não lhe trará nenhuma despesa ou risco. As informações recolhidas serão efetuadas através de uma pequena entrevista que deverá ser gravado para permitir uma melhor compreensão dos factos. A sua participação neste estudo é voluntária e pode retirar-se a qualquer altura, ou recusar participar, sem que tal facto tenha consequências para si.

Confidencialidade e anonimato: Qualquer informação será confidencial e não será revelada a terceiros. Os dados recolhidos serão de uso exclusivo para o presente estudo.

Identificação do/a investigador/a e da pessoa que pede o consentimento, se for diferente:

Raquel Maria do Espírito Santo Faria, estudante na Faculdade de Engenharia da Universidade do Porto.

Assinatura(s):

~~~~~

Declaro ter lido e compreendido este documento, bem como as informações verbais que me foram fornecidas pela/s pessoas/s que acima assina/m e que considero suficientes. Foi-me garantida a possibilidade de, em qualquer altura, recusar participar neste estudo sem qualquer tipo de consequências. Desta forma, aceito participar neste estudo e permito a utilização dos dados que de forma voluntária forneço, confiando em que apenas serão utilizados para esta investigação e nas garantias de confidencialidade e anonimato que me são dadas pelo/a investigador/a.

Nome: \_\_\_\_\_

Assinatura:

Data: \_\_ / \_\_ / \_\_\_\_

**ESTE DOCUMENTO, COMPOSTO DE 1 PÁGINA/S, É FEITO EM DUPLICADO:**

**UMA VIA PARA O/A INVESTIGADOR/A, OUTRA PARA A PESSOA QUE CONSENTE**

## APPENDIX C: Codes from NVIVO

| <b>Code:</b>                                         | <b>Description:</b>                                                                                                                                                                                                                              |
|------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Benefits</b>                                      | <b>User perceptions of benefits associated to AI</b>                                                                                                                                                                                             |
| Accessibility                                        | AI tools are easily accessible in terms of learning and understanding and use                                                                                                                                                                    |
| Future Advancements                                  | AI tools have the capacity to boost innovation and promote well-being in society                                                                                                                                                                 |
| Decision-Making                                      | AI tools helps users make easier and better decisions.                                                                                                                                                                                           |
| Efficiency                                           | AI tools can accomplish tasks with minimal wasted time and effort                                                                                                                                                                                |
| Personalization                                      | AI tools have the capacity to provide personalized solutions and experiences                                                                                                                                                                     |
| Reliability                                          | AI tools can be relied on to perform tasks                                                                                                                                                                                                       |
| <b>Concerns or Risks</b>                             | <b>User perceptions of concerns or risks associated to AI</b>                                                                                                                                                                                    |
| Concern about monopolistic power                     | Concern over how there are a small set of dominant companies and people in this field that can ultimately control the models.                                                                                                                    |
| Concern for biased algorithms                        | Concern for unfair or discriminatory algorithms that might treat them unfairly.                                                                                                                                                                  |
| Concern for Ethical Responsibility or Accountability | Concern over the grey area that AI provides ethically. Experts are concerned about accountability to when something goes wrong.                                                                                                                  |
| Concern for Misinformation or Biased information     | Concern about receiving inaccurate information, as well as information that favoured one side and did not state all the facts.                                                                                                                   |
| Concern for the future                               | Concern on how AI algorithms may impact our future in terms of control and advancements                                                                                                                                                          |
| Concern for transparency                             | Concern about how these models lack clear guidelines about consumers personal information and the decision criteria used by the models.                                                                                                          |
| Concern of not exposing personal information         | Concern over not exposing or disclosing personal information.                                                                                                                                                                                    |
| Concern of use or collection of Data                 | Concern on how their data was collected, used and stored by companies that they provided their information to.                                                                                                                                   |
| Concern over robot autonomy                          | Concern about AI performing automatic decisions without human validation.                                                                                                                                                                        |
| Concern over company actions                         | Concern for companies practices, such as consumer exploitation, purposely disseminating misleading content to generate attention and prioritizing competition in the market rather than safeguarding consumers in terms of privacy and security. |
| Concern over Data Ownership                          | Concern about who owned their personal data. Users were generally unaware as to who has access to their personal information.                                                                                                                    |
| Concern over robustness of systems                   | Concern about the reliability, resilience and stability of AI systems, that can produce errors, biases and unexpected behaviours                                                                                                                 |
| Concern over User Unawareness                        | Concern about how consumers were not aware of how the systems function and results are obtained from these algorithms.                                                                                                                           |
| Concerns over safety                                 | Concern on how these algorithms might have the potential to harm individuals or how they can contribute to unsafe outcomes.                                                                                                                      |
| Fear of being listened                               | Fear that someone, or something may be listening to them through electronic devices.                                                                                                                                                             |
| Fear of being watched                                | Fear that someone, or something may be watching them through electronic devices.                                                                                                                                                                 |
| Risk of Information breach or leak or hacked         | Concern over the risk of their personal information being breached, leaked or hacked.                                                                                                                                                            |
| <b>Disadvantages</b>                                 | <b>User perceptions of disadvantages associated to AI, that result in challenges for companies</b>                                                                                                                                               |
| Anthropomorphism                                     | Users might attribute human-like features to AI tools that might lead to unrealistic expectations and perceptions                                                                                                                                |
| Being misguided or restricted                        | AI tools might make users feels restricted or misguided by the information they provide                                                                                                                                                          |
| Default options                                      | AI tools may come with default options for users, which are difficult to understand and handle according to their preferences                                                                                                                    |
| Digital Footprint                                    | Users are not aware that all their information that is provided online, stays online, leaving a digital footprint                                                                                                                                |
| Functionality                                        | AI tools may not perform according to user expectations                                                                                                                                                                                          |
| Loss of Critical Thinking                            | AI tools promote that users rely on their systems, leading to a loss of critical thinking                                                                                                                                                        |
| No Human touch or interaction                        | AI tools are not capable of interacting as well as a human to human interaction                                                                                                                                                                  |
| Not in Control                                       | AI tools can overpower users and make them feel like they are not in control                                                                                                                                                                     |

| <b>Code:</b>                      | <b>Description:</b>                                                                                  |
|-----------------------------------|------------------------------------------------------------------------------------------------------|
| Trade off                         | For users to obtain any benefits from AI tools, there is always a trade-off of information.          |
| <b>Machine Learning artifacts</b> | <b>Machine Learning artifacts mentioned or used by the interviewees</b>                              |
| Chatbots                          | Any Machine learning artifact that has a chatbot feature                                             |
| Predictive analytics              | Any Machine learning artifact that has predictive analytics feature                                  |
| Recommender systems               | Any Machine learning artifact that has a recommender system feature                                  |
| Virtual Assistants                | Any Machine learning artifact that has a virtual assistant feature                                   |
| <b>Sensitive data</b>             | <b>Sensitive data relates to information users are not willing to provide to AI Systems</b>          |
| Asset information                 | Sensitive data related to physical assets such as properties, vehicles or other physical possessions |
| Bank details                      | Sensitive data related to bank information, such as card details and financial balance               |
| Family member information         | Sensitive data related to family members, such as knowing their birthday and how many children.      |
| Location                          | Sensitive data related to applications knowing the exact location at all times of the user           |
| Medical information               | Sensitive data related to medical information, such as medical records                               |
| Personal Address                  | Sensitive data related to addresses, such as home or work addresses.                                 |
| Phone number                      | Sensitive data related to having the users phone number                                              |
| Photos                            | Sensitive data related to images or videos                                                           |

