Ethical implications and privacy concerns for users of AI in Marketing.

By

Erika M. Hettiaratchy

(ID#: 20229532)

Submitted in partial fulfillment of the

Requirements for the degree of

Bachelor of Arts in

The School of Digital Business and Innovation

Thesis Supervisor: Professor Yee Heng Tan

Tokyo International University

2024

© 2024

Erika M. Hettiaratchy ALL RIGHTS RESERVED

Acknowledgments

I would like to express my deepest gratitude to Professor Yee Heng Tan for his unwavering support and guidance throughout this journey. His assistance was invaluable, from answering countless questions to helping me navigate the complexities of my thesis and coding challenges.

I am also profoundly thankful to my friends and family, whose unwavering love and encouragement over the past four years have been my source of strength during challenging times. Without them, this accomplishment would not have been possible.

Furthermore, I extend my heartfelt appreciation to the university and all my professors for their exceptional teaching methods and belief in me throughout my academic journey. Your dedication and commitment to my education have been instrumental in my success.

Thank you all for being a part of this incredible journey.

Ethical implications and privacy concerns for users of AI in Marketing.

By

Erika M. Hettiaratchy

ABSTRACT

In today's digital era, consumers rarely read the terms and service agreements. This paper examines the ethical implications and privacy concerns for users of AI in Marketing, classifies AI types, explores its effects on marketing, and analyzes case studies of AI misuse. While existing literature states consumers value their privacy, they often struggle to read Terms of Services (TOS) hindering informed decisions. This study addresses this issue by employing abstractive and extractive summarization techniques to simplify lengthy TOS documents. Enhancing summarization convenience may encourage more readership of TOS. By reading this paper, consumers will gain insights into the ethical and privacy aspects of AI-driven marketing, enabling proactive privacy safeguarding and informed decision-making.

Keywords: Ethics, Privacy, Users, Consumers, AI, Marketing

TABLE OF CONTENTS

Acknowledgments	3
ABSTRACT	4
Introduction	<i>€</i>
Research Question	7
Background	8
Analytical Framework/ Hypothesis/ Argument	9
Methodology	10
Potential constraints/challenges	12
Expected Conclusions	12
Significance of Study	13
Literature Review	14
Theme 1: Applications of AI and Effects on Marketing	14
The Benefits and Disadvantages of AI for Data Collection and Personalization	14
AI and Customer Engagement	16
Implications for B2B and B2C	18
Theme 2:Ethical challenges of AI in Marketing and privacy	20
AI may possess Biases	22
Theme 3: Impacts on consumer privacy due to unethical marketing practices	23
The Privacy Risks of AI Data collection	23
Proposed solutions of AI Data collection	25
Research Design	26
Extractive text summarization.	27
Abstractive Text Summarization.	36
Discussion	42
Conclusion	44
References	46
Appendices	51
Appendix A	51
Appendix B.	53

Introduction

Since the advent of Artificial Intelligence, concerns among consumers about maintaining their user privacy have been widespread. One such concern is the potential for companies to profit from or exploit sensitive customer information by selling it or using it to train AI models. This paper aims to highlight the importance of maintaining user privacy by answering the question "What are the ethical implications of using AI in marketing for consumers, and how can these be addressed to ensure that user privacy is not violated?"

This review is significant as AI has the potential to revolutionize the field of marketing, understanding and addressing the ethical implications of AI allows us to prevent companies from ruining user trust, and treating users as the product, it also allows us to have a competitive edge and improve AI literacy. AI has various applications in computing, data collection, personalization, engaging customers, and making processes more efficient for businesses and consumers(Vlačić et al., 2021; Chintalapati et al., 2022). As the field of AI continues to grow and evolve we as users need to understand the impacts of our choices so that we can make informed choices.

The purpose of this review is to explore existing literature and understand the ethical challenges surrounding the use of AI in marketing and its impact on consumer privacy. It also aims to provide insights into potential solutions to address these challenges. To achieve this, the review will follow a logical order, first focusing on the applications of AI and its effects on marketing, then exploring the ethical challenges of AI in marketing and privacy, and finally examining the impacts on consumer privacy due to unethical marketing practices.

This review will employ Davenport et al. (2020)'s theoretical Framework and classification system, "A Multidimensional framework to anticipate AI evolution: levels of intelligence, task type, AI in robots by Davenport, Guha, Grewal, and Bressgott". By utilizing this framework, we can gain a comprehensive understanding of the multidimensional aspects of AI's impact on marketing and consumer privacy. Kim et al. (2019)'s theoretical framework, "The privacy calculus theory framework by Kim, Park, and Ahn" will also be used to examine the factors influencing individuals' behavioral reactions toward information disclosure.

A significant gap in the existing literature is the lack of a user-friendly method for consumers to read and understand the terms and conditions associated with AI-powered services. To address this, this paper proposes a solution in the form of a program that summarizes terms of services, making it easier for users to understand what they're signing up for. The ultimate goal is to emphasize the significance of data privacy, increase customer trust, and potentially drive policy makers to change existing laws to protect consumers. However, it is important to acknowledge the limitations of this project, including time constraints, the efficacy of the program, and the validity of my input data from the dataset. Given the rapidly evolving nature of AI, it is important to note that some information mentioned in this paper may be outdated in a few years.

Research Question

What are the ethical implications of using AI in marketing for consumers, and how can these be addressed to ensure that user privacy is not violated?

Ethical and Privacy concerns for Al users in Marketing Background

AI, short for Artificial Intelligence, refers to an intelligent machine that possesses the capacity to learn, reason, self-correct, and even exhibit creativity. It does this by collecting, cleansing, modeling, and deploying vast quantities of user-generated data, with the aim of benefiting both consumers and businesses (Blier et al., 2019; Chintalapati et al., 2022; Grewal et al., 2021). In contrast, the Internet of Things (IoT) connects and automates processes while utilizing sensors and effectors to perceive and respond to its environment. For the purposes of this paper, the terms AI and IoT will be used interchangeably.

Voice Based Digital Assistants (VBDA) like Siri and Google Assistant, are examples of AI that collect and process data to improve consumer efficiency (Vimalkumar et al., 2021).

This topic is significant as AI has the potential to transform marketing strategies (Vlačić et al., 2021). This paper classifies AI into three types based on intelligence, task type, andpresence of effectors, using Davenport et al.'s (2020) method. The privacy calculus theory framework by Kim et al. (2019) is also used to understand factors influencing individuals' reactions to disclosing information. Consumer privacy involves managing others' presence during transactions and controlling information sharing.

This paper will use the definition of privacy by Goodwin (1991, as cited in Foxman et al., 1993) Consumer privacy as the ability of individuals to manage the presence of others during transactions, and regulating the sharing of information related to those transactions or activities with individuals who were not present.

Analytical Framework/ Hypothesis/ Argument

My thesis argues that while the use of AI in marketing raises ethical concerns as discussed by

Castelo et al. (2018) regarding user privacy these implications can hypothetically be addressed through the implementation of regulations, transparency in data collection practices, and the development of privacy-enhancing technologies. I will use Kim et al. (2019)'s theoretical framework, "The privacy calculus theory framework by Kim, Park, and Ahn" which will allow me to compare the factors influencing individuals' behavioral reactions toward information disclosure and allow me to determine what factors affect the tradeoff between perceived privacy risk and benefit affect consumer willingness to provide private information in various contexts. In addition to this Davenport et al. (2020)'s Multidimensional framework to anticipate AI evolution: levels of intelligence, task type, AI in robots, will also be used to classify AI and their impacts on users. The significance of this claim lies in the growing prevalence of AI in marketing and the increasing reliance on user data for personalized advertising and consumer profiling while contributing to the ongoing conversation surrounding the responsible and ethical use of AI in the marketing industry as discussed by Chintalapati et al. (2022) and Grewal et al. (2021).

Davenport et al. (2020) conducted a study investigating the influence of AI on marketing and consumer behaviors. It explores AI implementation methods and categorizes them by level of intelligence, task type, and presence of effectors and emphasizes the need for further research in areas such as marketing strategies, customer behaviors, data privacy, bias and ethics. As my topic is quite new it has few seminal articles, Davenport is an established researcher and this paper in particular is cited by many others for AI classification.

Summaries of research papers cited in this article are also provided, thus making it easier to grasp what factors the author considered before they reached their conclusion. This comprehensive analysis provides insights into the impact of AI on consumer

information disclosure, supporting future research in the field. This paper provides a valuable insight and foundation for understanding the implications of AI in marketing strategies and consumer behaviors. It also addresses the importance of data privacy and other ethical considerations, which aligns with my focus on consumer privacy. And so this text acts as a basis to understand the importance of AI and the ways in which consumers may benefit while also being a device to glean important sources from based on their summary, and finally, a way to highlight the advantages and disadvantages of AI in certain fields.

Methodology

This paper aims to answer the question, "What are the ethical implications of using AI in marketing for consumers, and how can these be addressed to ensure that user privacy is not violated?" to determine the effects of AI on consumers. Do consumers fully understand the implications of AI? Are users informed about how their data can be used for marketing purposes? And finally, Are users aware of the potential risks associated with AI and its interpretive abilities? This paper aims to answer these questions and create a solution in the form of a summarization program that simplifies terms and services for consumers.

The framework used is "A Multidimensional framework to anticipate AI evolution: levels of intelligence, task type, AI in robots by Davenport, Guha, Grewal, and Bressgott" This is utilized in classifying AI types and their impacts on users. In addition, the framework by Putoni et al. (2021), which takes a customer-centric approach to AI, will be applied to understand the effects of AI on marketing and its implications for consumers.

This paper's target audience is primarily social media users ranging from school-age to adults. A publicly available terms and services database will also train my dataset. This is

at(https://www.kaggle.com/datasets/sonu1607/tosdr-terms-of-service-corpus?resource=download). This work was inspired by Roy et al. 's (2012) website (https://tosdr.org), which classifies the terms of services of varying companies by using user collaboration to summarize various websites' TOS(Terms Of Services). I aim to enhance this process through automation, making it more efficient. Two approaches are used to assess AI: a method to classify AI based on Davenport et al. (2020) and a Customer-centric approach to AI based on Putoni et al. (2021). This paper's structure was inspired by Berton et al. (2020).

To collect data, I employed existing surveys from published resource papers within my literature review. I will also make the

extractive summarization (https://www.kaggle.com/code/ericahh/extractive-summarization)
and abstractive summarization

(https://www.kaggle.com/code/ericahh/abstractive-summarization-with-bart-tosdr)
code freely available on Kaggle so that those who are interested can use the code to
summarize TOS that they find difficult to read (Hettiaratchy, 2024). The outcome of these
summarization techniques will determine the effectiveness of my proposed method. By
automatically summarizing terms and services using a database of keywords and phrases
users may be more likely to read the critical aspects of TOS before agreeing to them. This
will improve the existing website by Roy et al. (2012) as reducing the reliance on user
submissions will allow for even lesser-known websites to be analyzed easily.

Potential constraints/challenges

The theoretical framework will serve to classify AI types and their impacts on

users. The summarization program will act as a technical tool to educate users on what they are agreeing to when accepting terms and services. This classification system is critical to my research with regards to classifying my primary data. In order to carry out procedures that will summarize terms and services, it is important for me to classify words by preprocessing them and for this I will also use the aforementioned dataset (Gupta, 2023) to train my model. My constraints and challenges are currently a need for more technical knowledge. I will ensure that I learn these skills by the proposed time frame to create a viable product.

Expected Conclusions

The research question for this study is "What are the ethical implications of using AI in marketing for consumers, and how can these be addressed to ensure that user privacy is not violated?" Based on the findings and analysis conducted several expected conclusions can be drawn.

This study emphasizes the importance of user privacy in the context of marketing and AI and aims to restate the significance of protecting user data. This study also highlights the need for policymakers to enforce methods that encourage consumers to be aware of how their data is being collected, stored, and utilized by companies. As demonstrated by the General Data Protection Regulation (Sun & Medaglia, 2019), by promoting transparency and accountability in data practices, consumers can make informed decisions about how their personal information is being used, ultimately empowering them to exercise greater control over their data.

Moreover, this research intends to contribute to existing knowledge by shedding light on the negative aspects of AI and its potential implications for user privacy. By raising awareness

among the general public, it is expected that consumers will become more cautious and informed about terms of service documents and AI. This increased awareness can lead to enhanced user experiences, as consumers will be better equipped to navigate the digital landscape and protect their privacy rights.

However, it is important to acknowledge the limitations of this research. One potential limitation is that it is assumed that users of the program to summarize terms of services, will be willing to invest the time and effort to learn how to protect their data and follow the practices as intended. Additionally, there is a possibility that the TOS in my dataset (Gupta, 2023) are out of date. To conclude, this study underscores the significance of user privacy in the realm of AI driven marketing. It advocates for the implementation of policies that prioritize transparency and user empowerment, aiming to protect individuals' data rights. While there may be limitations to the research, such as the accuracy of my dataset (Gupta, 2023) and other constraints, it is believed that these findings will contribute to the ongoing conversation of the ethical implications and privacy concerns of AI in marketing for users. By addressing these concerns, this research aims to create a more ethically responsible and privacy-conscious marketing landscape.

Significance of Study

The implications of AI in marketing for consumers and their privacy is important because it is one of the most widely debated topics regarding AI in the present day. As AI is an emerging field, it is crucial in terms of today's concern over privacy. Ultimately, what is at stake here is consumers freedom of speech and choices. Consumers that are targeted with ads based on their demographics and interests do not overlap with consumers with different interests, this may lead to a growing divide between demographics. These findings have

important implications for the broader domain of internet safety and freedom of choice.

Consumers have a right to know which information companies collect from them and the purpose behind its collection.

Literature Review

Theme 1: Applications of AI and Effects on Marketing

The Benefits and Disadvantages of AI for Data Collection and Personalization

One of the most important aspects of AI is its ability to collect large amounts of consumer data, segment them into different categories, and output specific information targeted at consumers based on their likes and dislikes; This is called Data collection and personalization. Chintalapati et al. (2022) emphasize that AI collects and processes important data to benefit consumers and businesses, allowing customized marketing techniques. Personalization increases the time spent by users on their platforms which in turn increases engagement and profits. Personalization is a common practice used in any social media website. For example, Instagram and Tiktok personalize their reels for each consumer to increase interactions and time spent on the app by consumers. Grewal et al. (2021) also highlights AI's role in real time data capture, processing, and providing useful outputs to consumers. This can be extremely helpful for consumers that rely on real time updates due to health concerns For example, AI can also contribute to consumer health monitoring, personalized updates, suggestions, and customer engagement. We observe that by collecting data and personalizing advertisements we can predict customer wants and needs, reduce search costs, and provide convenience (Grewal et al., 2021,2020).

In this way AI has completely changed the marketing landscape compared to the era

of newspapers and billboard marketing. Huang and Rust (2021), state that the classification and application of AI had significant implications for marketing research, strategy, and action. This is supported by Vlačić et al. (2021) who states that AI has the potential to revolutionize marketing strategies and channels by being able to understand customers and through online targeting, recognizing niche markets and determining the potential of certain markets leading to more profitable business models. By predicting trends and niches before competitors, marketers can increase their profits by being a leading brand. This would also benefit consumers who would have their specific needs met much faster. These aspects are also important for marketing agencies that depend on click-through rates and engagement. Consumers' intentions to use IoT services depend on perceived benefits, attitude, and privacy Hsu et al. (1993). However, too much data collection could also lead to distrust. When it comes to healthcare services involving sensitive information, consumers hesitate to grant companies access to such data, despite the potential trade-off of receiving less personalized experiences(Kim et al., 2019). Distrust then leads to consumers losing loyalty, as shown by Foxman et al. (1993)'s discovery that the disadvantages of unrestricted utilization of consumer information outweigh the advantages. This correlates with Hsu et al. (1993)'s findings that repeat transactions are unlikely to occur if consumers are aware of privacy infringements, as users are primarily concerned about unauthorized access and usage of their personal information rather than the amount of services they can receive in exchange (Hsu et al., 1993). In conclusion, AI's ability to collect and personalize consumer data has revolutionized marketing. It enables targeted marketing, increases engagement and profits, and benefits consumers with real-time updates. Al's application in marketing research and strategy allows for understanding customers, predicting trends, and identifying niche

markets. However, excessive data collection can lead to consumer distrust and privacy concerns, impacting loyalty. Striking a balance between data collection and privacy is vital for maintaining trust in the AI-driven marketing landscape.

AI and Customer Engagement

AI can also be used for customer-centric roles. Vimalkumar et al. (2021) found that voice-based digital assistants (VBDA) enable consumers to interact more intuitively with their devices. VBDA utilizes voice recognition technology and natural language processing to understand and respond to user commands and inquiries. They can perform various tasks such as answering questions, providing recommendations, setting reminders, and more. VBDAs like Alexa, Siri, Google Assistant, and Cortana Vimalkumar et al. (2021) are commonly integrated into smartphones, smart speakers, and other devices. VBDA has been a game changer for those with disabilities as they are able to use these devices to buy necessities allowing them to become more independent. In addition to VBDA, AI plays a crucial role in marketing. AI automation, data insights, and customer engagement are key aspects of utilizing AI in marketing (Davenport et al., 2020). The use of chatbots, virtual assistants and VBDA allow for companies to enter consumer homes and understand their needs, by paying attention to keywords that are spoken and providing efficient responses in the form of advertisements. Marketers can enhance overall customer experiences. However, This may pose a risk as most users are unaware of their data collection when using VBA (Pitardi et al., 2021). AI can also improve relationship marketing; this involves factors such as trust and power dynamics which are crucial in improving marketing campaigns (Grewal et al., 2021). Relationship marketing focuses on fostering strong connections and trust with customers rather than focusing on individual transactions. Pitardi et al. (2021) et al supports

this argument with his findings that consumers interact with Voice based digital assistants as social entities and develop deep connections with them. This is known as a parasocial relationship and therefore users have increased trust in VBDA's. While this may sound dangerous he eases our worries by stating that VBDAs are often used for home automation and so they are not perceived as violating privacy. (Pitardi et al., 2021). Fostering parasocial relationships increases customer loyalty and satisfaction by fulfilling their needs and preferences consistently. Without AI, this task would require an unattainable amount of manpower.

Many studies have been conducted on the effects of AI in the digital marketing and e-commerce fields. A few of these are Integrated digital marketing, content marketing, experiential marketing, marketing operations, and market research are some of the themes highlighted by Blier et al. (2019). Integrated digital marketing is the use of marketing tactics such as SEO, social media marketing, Email marketing, etc Blier et al. (2019). Integrated marketing aims to create a cohesive and engaging experience for the customer. Blier et al. (2019) also explains content marketing, which, in contrast, focuses on creating and distributing valuable and relevant content to attract, engage, and retain a target audience. This can be in the form of blog posts, videos, infographics, etc. This marketing strategy's purpose is to build brand awareness, increase customer engagement and increase conversions.

In summary, AI has opened up new possibilities for customer-centric roles, allowing for more intuitive interactions, personalized marketing strategies, and enhanced relationship building. As AI continues to advance, its application in customer-centric roles will likely expand further, benefiting both businesses and consumers alike.

Implications for B2B and B2C

The implications of AI in marketing for both B2B (business-to-business) and B2C (business-to-consumer) sectors are significant for small companies and large companies alike. By using marketing techniques that are outdated, companies may not be able to compete with current methods done with AI such as personalization and niche identification. Several studies conducted by Grewal et al. (2021) and Chintalapati et al. (2022) have provided detailed insights into these effects of AI on B2B and B2C companies. Vlačić et al. (2021) analyzed 164 academic articles and uses content analysis and statistical techniques to analyze the literature and concluded that In both B2B and B2C companies, AI has the potential to revolutionize marketing strategies and channels by being able to understand customers, utilizing online targeting semantic recognition which is when customer databases are created and analyzed for useful insights. By utilizing AI in these ways companies can use AI to enhance efficiency and target greater numbers of customers. AI offers opportunities to enhance efficiency and efficacy. AI also has the ability to recognize niche markets and determine the potential of certain markets leading to more profitable business models (Vlačić et al., 2021). AI-powered tools can automate various marketing tasks, such as lead generation, customer segmentation, and personalized messaging, allowing businesses to streamline operations and allocate resources more effectively. AI could also be used in service industries to lead to higher production (Grewal et al., 2021). Such as automated restaurants and using VBDA as technical support. AI can automate repetitive tasks and conduct simple transactions (Vlačić et al., 2021). This could be useful in unmanned stores.

AI can also assist in predictive analytics, helping B2B companies identify patterns and trends, thereby facilitating better strategic decision-making Grewal et al. (2021).

Additionally, It can contribute to improving customer relationship management and enhancing customer experiences through personalized interactions and tailored solutions

Chintalapati et al. (2022).

AI-powered algorithms and machine learning techniques enable businesses to collect and analyze vast amounts of consumer data leading to valuable insights for understanding consumer behavior, preferences, and purchase patterns Chintalapati et al. (2022). This information can be utilized to develop targeted marketing campaigns, personalized recommendations, and customized experiences that resonate with individual consumers. For instance, Kim et al. (1993) found Recommendation systems help users find information efficiently and reduce the need for searching and selecting. However, these systems also raise privacy concerns as private information about users must be inferred to display these recommendations.

In conclusion, AI has significant implications for marketing in both B2B and B2C sectors. It enables companies to personalize marketing efforts, identify niche markets, automate tasks, and analyze consumer data for valuable insights. AI-powered strategies enhance efficiency, facilitate better decision-making, and improve customer experiences. However, privacy concerns need to be addressed when utilizing AI for personalized recommendations. Overall, AI revolutionizes marketing strategies, allowing businesses to stay competitive and achieve long-term success.

Theme 2:Ethical challenges of AI in Marketing and privacy

The implementation of AI in marketing introduces a range of ethical implications for

consumers. These implications include issues of consumer distrust, power imbalances, biases, ethical considerations, data privacy, and the limitations of AI in understanding human emotions (Chintalapati et al., 2022; Grewal et al., 2021; Castelo et al., 2018; Luo et al., 2019). These challenges require careful consideration to ensure user privacy is not violated and ethical practices are upheld in AI-driven marketing strategies.

In the context of IoT services, Hsu et al. (1993) found that consumers' intentions to use IoT (Internet of Things) services depend on perceived benefits, attitude, and information privacy. This indicates that consumers are willing to trade their data in exchange for benefits. However, an important factor is whether consumers are aware of the potential risks associated with the exchange of their personal information for temporary benefits. Antón et al. (2010) had different results when they conducted a survey in 2008 which showed that individuals' primary privacy concerns of information transfer, storage and awareness remain unchanged since 2002, and instead increased in some areas.

It has been observed that users with prior experience in Internet-related services are more inclined to accept and adopt IoT services (Hsu et al., 1993). This may explain why current research indicates that users are becoming more lax regarding their privacy in digital environments. As users continue to utilize AI and IoT in their daily lives, they may be more willing to share sensitive information. However, this may create a negative feedback loop where users perceive their privacy as unimportant, stop educating themselves on data collection practices, and consequently provide more data. Foxman et al. (1993) argue that privacy depends on whether consumers are aware of their data being collected, their privacy rights, and which party controls their data. If consumers are unaware of their data being collected, it can be argued that they have no privacy as they are unable to give informed

Consumer distrust is a powerful argument associated with AI in marketing.

Chintalapati et al. (2022) note that consumers are wary of AI due to a power imbalance and distrust. The power imbalance is due to users being unaware of which data AI collects and processes to make assumptions about user preferences and usage. This sentiment is supported by Grewal et al. (2021), who highlight the untrustworthiness of AI and emphasize the need to address the power imbalance between consumers and AI systems. Furthermore, Castelo et al. (2018) and Grewal et al. (2021) discuss how AI can introduce biases and ethical challenges by grouping users by race, age, and interests and predicting their actions and interests, further exacerbating consumer distrust.

Ensuring ethical processes, data usage, and generation while maintaining consumer privacy presents ongoing challenges and potential legal issues (Sun & Medaglia, 2019). The ethical implications of using AI require organizations to prioritize transparency and accountability. By providing clear information about data collection practices, processing, and utilization, companies can help build consumer trust. Another way to build customer trust on a larger scale is to implement legal frameworks and regulations that protect consumer privacy and address ethical concerns, such as the General Data Protection Regulation (GDPR), which has already been shown to be successful in Europe (Sun & Medaglia, 2019).

Additionally, AI's limited ability to understand human emotions poses another ethical challenge in marketing, as argued by Luo et al. (2019). They highlight negative consumer responses to interacting with bots and the need to address AI's shortcomings in comprehending and responding appropriately to human emotions. To mitigate this issue,

organizations should invest in advancing AI systems' capabilities through natural language processing, sentiment analysis, and machine learning, enabling a more accurate interpretation and response to human emotions.

AI may possess Biases

Furthermore, the potential for biases in AI algorithms raises concerns regarding discrimination against vulnerable consumer groups. Davenport et al. (2020) and Mogaji et al. (2020) agree that AI could develop biases leading to discrimination against vulnerable consumer groups; this may lead to ethical dilemmas and may pose legal issues. However, Putoni et al. (2021) argue that AI may level the playing field and mitigate discrimination while having the potential to address pressing societal problems if designed and implemented appropriately.

In summary, the ethical implications of using AI in marketing for consumers encompass challenges such as distrust, power imbalances, biases, ethical considerations, and limitations in understanding human emotions are major concerns that should be addressed. Protecting consumer privacy and addressing concerns is crucial to building trust as brought up by Chintalapati et al.(2022) and Grewal et al.(2021). Organizations must prioritize transparency, accountability, and responsible data usage to address these concerns and ensure user privacy is not violated. Additionally, enhancing AI systems' understanding of human emotions and mitigating biases are crucial steps toward ethical AI implementation in marketing. Adherence to legal frameworks and regulations also plays a vital role in protecting user privacy and addressing ethical implications. By navigating these challenges and proactively addressing them, organizations can foster trust, maintain user privacy, and promote the ethical use of AI in marketing.

Theme 3: Impacts on consumer privacy due to unethical marketing practices

The Privacy Risks of AI Data collection

After conducting this research, it was evident that all sources agree that AI data collection poses a significant threat to privacy as supported by both Chintalapati et al. (2022) and Mogaji et al. (2020). Consumers express concerns regarding their privacy when using AI due to unintentional data collection and potential misuse of sensitive information; these concerns have been highlighted by Chintalapati et al. (2022), Gardiner (2018), and Vimalkumar et al. (2021).

The Vulnerability of devices also poses another threat Zhou et al. (1993) state the interdependence of IoT could enable attackers to target smart devices to perform actions that will lead to a security breach. For example, a smart AC could be turned off to increase temperature and open windows. The diverse nature of IoT systems could have weak authorization allowing attackers to find vulnerabilities within systems leading to data breaches (Zhou et al., 1993). IoT devices lack established defenses as they are quite new.

Users may also be unaware of potential risks or complacent, Based on their research, Kim et al. (2019) discovered that consumers often disregard privacy risks to receive more personalized services. However, another author Kim et al. (1993) found that perceived benefits, convenience, and vulnerability did not significantly affect the intention to disclose personal information. This leads to mixed results in this section.

Unintentional data collection of health data is a pretty sensitive topic. Wearable devices, smart meters, etc. collect biological information, this intimacy may lead to privacy concerns, as individuals utilizing Internet of Things (IoT) services anticipate the secure

Ethical and Privacy concerns for Al users in Marketing retention and non-disclosure of their personal information to external entities (Kim et al., 2019; Zhang et al., 1993; Zhou et al., 1993).

One of the key challenges arises from AI's operant nature(it's ability to collect information and make educated guesses) can lead to the inference of data that users may not even know about themselves. Collecting sensitive information without user consent compounds privacy concerns associated with AI. Vimalkumar et al. (2021)

provide examples where the potential collection of data, such as location history, purchase history, calendars, contacts, and voice queries, may be collected without the user's consent(Vimalkumar et al., 2021).

In conclusion, the research clearly indicates that AI data collection poses a significant threat to privacy. Multiple sources, including Chintalapati et al. (2022) and Mogaji et al. (2020), support this view. Consumers express concerns about unintentional data collection and potential misuse of sensitive information. The vulnerability of IoT devices, as noted by Zhou et al. (1993), further amplifies privacy risks. Mixed findings regarding user behavior and awareness highlight the need for further investigation. Health data collection through wearable devices and smart meters raises specific privacy concerns. The operant nature of AI, allowing for inference of data without user knowledge, compounds privacy worries. Safeguarding privacy while reaping the benefits of AI requires attention from researchers, policymakers, and industry stakeholders.

Proposed solutions of AI Data collection

In response to these challenges, politicians and implementers such as programmers and big corporations should recognize the importance of maintaining privacy. Maintaining privacy is a major concern for implementers, highlights Mogaji et al. (2020). Protecting user

privacy builds trust and reduces the likelihood of future legal issues. Companies attempt to address the issue of AI inferring data by anonymizing data to negate these risks. However, users should still be informed of potential risks and be willing to share their data as emphasized by Grewal et al. (2020), and Putoni et al. (2021).

It is also crucial to determine which customers are willing to share their privacy for certain benefits and services. Presently, this is done through terms and services. However, this method could be more effective as this process is lengthy and complicated. Grewal et al. (2021) state that expecting consumers to understand and choose to share their information repeatedly may not be feasible. To address these privacy concerns, transparency, and informed consent are crucial. AI can potentially automate this process (Grewal et al., 2021), providing users with greater visibility into data collection and usage. Putoni et al. (2021) state that it is essential to balance utilizing data for marketing purposes and respecting consumer privacy.

In conclusion, addressing privacy challenges requires recognizing its importance and protecting user privacy to build trust and prevent legal issues. Anonymizing data and informing users of risks are necessary steps. To determine willing data sharers, a more effective and automated approach is needed, ensuring transparency and informed consent. Balancing data usage and consumer privacy is crucial.

Research Design

This study aims to analyze the ethical implications and privacy concerns associated with AI in marketing by focusing on how terms and services agreements can be made more comprehensible to users. To implement the aforementioned proposed solutions. We will employ both extractive and abstractive summarization techniques to distill these complex

documents into user-friendly summaries, thus empowering users with a clearer understanding of the information they are consenting to share. To evaluate the effectiveness of these summarization methods the percentage summarization will be calculated and the run time and ease of access to users will be considered. Additionally, in final processes the accuracy of the summaries could be assessed by comparing them against the original documents to ensure no critical information is lost.

Ethical considerations are integral to this study. Transparency in the summarization process is crucial, therefore the sources of code will be made readily available, ensuring users are aware of the methods used. Techniques to minimize biases in the summarization models will be applied to ensure fairness. In addition to this, the same dataset will be used for both summarization techniques. By employing these methodologies, this thesis aims to contribute to more ethical AI practices in marketing and enhance user awareness and control over their personal data by providing the means for users to summarize TOS data themselves.

Extractive text summarization

According to Zhao (2018), extractive text summarization combines words within the source and tries to create a sentence with meaning. This way it creates a summarization of the document while preserving key information. There are multiple techniques that can be used for this method such as ranking words and sentences in order of importance and weighing keywords. This method requires preprocessing; which is the

removal of elements that are not text such as punctuation and stop words. It is then tokenized and classified as a noun, verb or adjective. The most relevant words are then ranked and concatenated to make a coherent sentence.

When conducting extractive text summarization, I followed the methods that Zhao (2018) used. The TextRank algorithm from Google's PageRank algorithm was used. This way the most relevant sentences were identified from the terms of services dataset (Gupta, 2023). The text files were then preprocessed to remove unnecessary elements like common stop words and advertisements thereby retaining only nouns and adjectives. The TextRank algorithm then ranks the sentences according to their importance, thereby capturing the main points of the document (Zhao, 2018). My keywords are set relating to privacy concerns namely, 'privacy', 'data', 'personal information', 'cookies', 'tracking', 'security', 'consent'. By changing these keywords the output of this model should change. A loop is also included that allows the user to run each file. For each file it prints out a summary and details about the length of the document (to view the full summary see Appendix A).

To demonstrate the code in action the Juicer_CookiePolicy text file from the TOS data set was used as a test to determine the following data in Fig 1.

Fig 1: Sample extractive summarization output for the Juicer CookiePolicy text file

```
Original length of the document: 4715 characters
Summarized length of the document: 4713 characters
Reduction in length: 0.04%
Run time for processing the file: 0.05 seconds

Summary of the document:
It is recommended that you leave on all cookies if you are not sure whether you need them or not in case they are used to provide a service that you use. How to disable CookiesIf you do not want information collected through the use of cookies when using your Juicer account or our website, there is a simple procedure in most browsers that allows you to automatically decline cookies or gives you the choice of declining or accepting the transfer to your computer of a particular cookie (or cookies) from a particular site. We may also use cookies to see how you interact with our offers, and we may use cookies or other files to understand your use of our websites....
```

The summarization is still lengthy and could be improved. To view the whole summarization please refer to Appendix A However the privacy concerns identified in the document are displayed and does appear to give us a coherent answer. From Fig 1 we can determine that the cookies in the Juicer website collect browser type, time spent on the digital Services, pages visited, etc. The user may find this data useful when considering accepting the TOS from this website.

After running the extractive summarization technique for each file and calculating the average reduction in length for all files, we determined that this method was 76.12% effective at condensing the TOS we provided (see Fig 2). We are also able to determine the averages for all the files in the ToSDR dataset (Gupta, 2023). As shown in Fig 2, after running 9496 files, the average reduction in length for all files was 76.12% and the average run time for each file was 0.303 seconds.

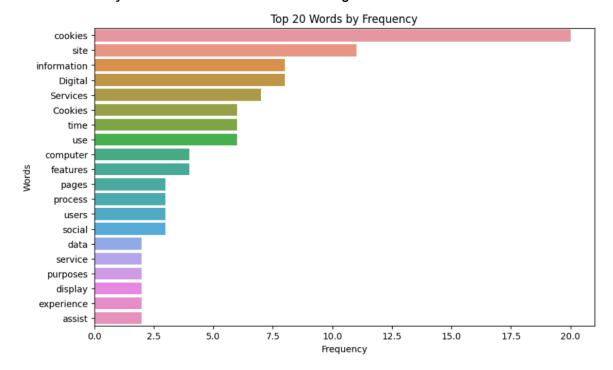
```
File: Sucuri_PrivacyPolicy.txt
Original length of the document: 19334 characters
Summarized length of the document: 3249 characters
Reduction in length: 83.20%
Run time: 263.98 milliseconds

Average reduction in length for all files: 76.12%
Average run time for all files: 303.02 milliseconds
```

Fig 2

Additional information can be gleaned through the extractive summarization technique being run on Kaggle. The graph in Figure 3 shows us that the privacy keywords that were used are mostly effective.

privacy_keywords = ['privacy', 'data', 'personal information', 'cookies', 'tracking', 'security',
'consent']



On the y axis certain privacy-related terms such as "cookies", "data", "features" align with the keywords that were entered. This could allow users to infer the most frequent words in this text file. The term "cookies" appears with the highest frequency of 20. It's capitalized form of "Cookies" is also mentioned as measuring in at 6 on this scale. This suggests that the use of cookies is a significant aspect of how the TOS of this website handles user data and tracks their behavior. This could be important for the user to understand as this could impact their privacy. Terms like "data," "service", purposes," "display," "experience," and "assist" have lower frequencies of 2. Which suggests that these terms are less frequent in this text file compared to "cookies". By having this graph we can improve upon our summarization by ensuring that high-frequency keywords are included in the summarization. If certain privacy keywords are referred to less it might indicate areas where the TOS documentation for this website is lacking. For example, if consent is not mentioned, the website may not take users' consent into consideration when collecting their data.

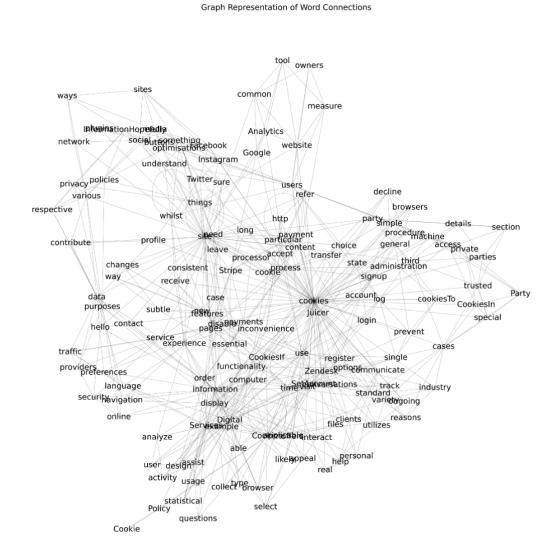
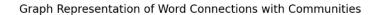


Figure 4

Another method of visualizing the data from our test file is through the use of word connections. We can see from the graph in Figure 4, that the main clusters revolve around the terms "cookies," "optimizations," "Services," "pages," "size," and "Interact.". This visualization was ran with a cluster (K) value of 50. Upon visualizing this code numerous times by entering multiple K values, we can see that these same 6 clusters are the most prominent. Therefore we can assume that these are the most connected and predominant words in this file.

However we can also see repetitions in this graph such as with cookies, cookiesTo, CookiesIf, etc. This could demonstrate that there are issues in the file itself or with the existing model, perhaps the existing model is case-sensitive and therefore unable to group similar words together if they are different cases.



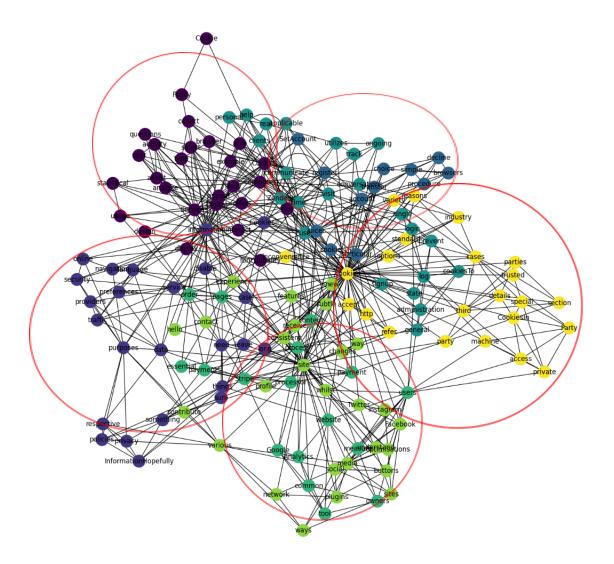


Figure 5

The graph was run again with the Force Atlas layout algorithm resulting in Figure 5. We can now determine the graph's community structure, showcased within the red circles.

These nodes are more closely connected and 5 clusters are shown. To interpret this data, when the model runs there are 5 different distinct groups that occur close to each other and so these are the most likely to have multiple connections and similarities ("Learning to read", 2022).

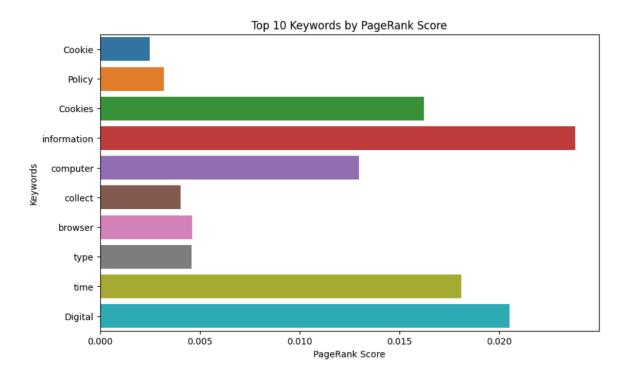


Figure 6

We can also determine the PageRank score, this is a formula that displays the keywords that are highly rated by Google's search engine (Varangouli, 2023). Here we see information, Digital, time and cookies are in the lead. The page rank scores scale from 0 to 10. 0 being the lowest quality pages and 10 being the highest quality. As we can see above, these individual keywords on their own have a low page rank score, this is as the page rank is calculated by considering the website and its links as a whole (Varangouli, 2023).

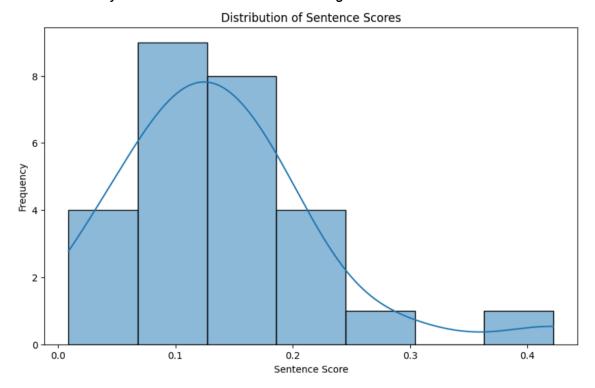


Figure 7

This histogram demonstrates the frequency distribution of sentence scores, The blue curve represents the probability density function (PDF), to visualize the distribution's shape. Most sentence scores are between 0.0 and 0.2. We also see that the distribution is right skewed showing that lower sentence scores are more common. Sentence scores display the importance of sentences. As this graph represents a higher concentration of sentences with lower scores this shows that the model is capturing less important information.

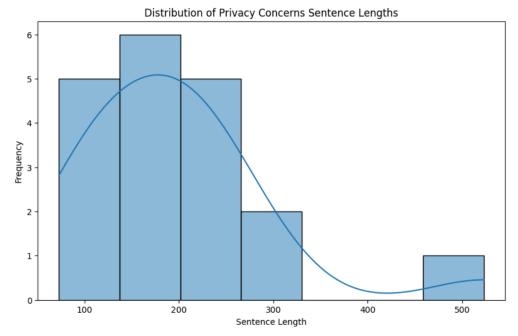


Figure 8

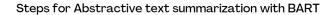
From the Distribution of privacy concerns sentence lengths graph, we can see the frequency distribution of the sentence lengths related to privacy corners. The PDF is also included here to help us visualize the graph. We can see that most sentences in the file related to privacy concerns have lengths between 100 and 200. However, there is a decrease in frequency for sentence lengths greater than 350 and smaller than 450. But there is a Frequency of 1 for sentence lengths between 450 and 550. This graph demonstrates whether the summarization model is addressing the privacy concerns aspect. As the privacy-related sentences are long, the tool is picking up these points. However further summarization might be necessary to highlight the most important aspects of the test file.

Abstractive Text Summarization

Abstractive text summarization, generates summaries that convey the core concepts and essential information of the given dataset (Gupta, 2023), in this case, the TOS dataset which is summarized in a more human-like, paraphrased manner (Rao, 2024). It is described as human-like as it can use words that are not in the text to convey its message effectively.

This paper will use BART, a transformer for abstractive text summarization. A Transformer is a model that uses an encoder-decoder structure to process inputs and produce outputs. The BART model or Bidirectional and Auto-Regressive Transformer is an improved model based on BERT and GPT models that are now outdated. The encoder generates embeddings for each token in the input text, while the decoder uses these embeddings to generate coherent and clean text sequences(Swain, 2023). This setup allows BART to handle a range of NLP tasks including text generation, translation, summarization, and paraphrasing thereby surpassing BERT and GPT-1 in performance(Swain, 2023).

When conducting the Abstractive summarization. The same set of terms and services agreements used for extractive summarization will be processed. Key concepts and ideas within the text will be identified using natural language processing (NLP) techniques, namely BART to create summarizations and further our analysis. These generated sentences will then be combined to create cohesive and concise summaries that are easier for users to understand(Zhao, 2018).



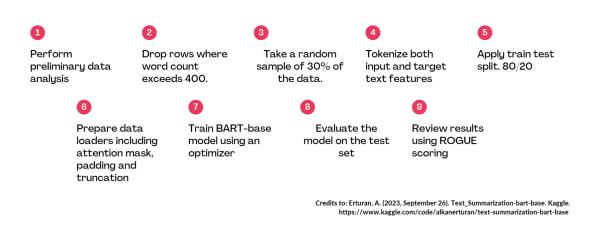


Figure 9 (Erturan, 2023)

By following Erturan's (2023) guide to text summarization with BART. A summary of each file was first created as this was missing from my training dataset (Gupta, 2023). A preliminary data analysis was then conducted followed by the tokenization process.

Tokenization is defined as a method to split up the text into smaller sections or "tokens". For this program, the length of the model was limited to 512 words. A random sample of 30% of the dataset's input and target text features was then tokenized. A train test split of 80/20 was used to train the BART model using an optimizer, which is an algorithm that adjusts the parameters of the model during training to minimize losses. The optimizer has 3 key functions, updating parameters, managing the learning rate of the model, and converting the model to produce the best solution without losing functions. The test set was then run and its prediction was evaluated using ROUGE (Recall-Oriented Understudy for Gisting Evaluation)

scoring. A ROUGE score is a measure to calculate the quality of a summarization model. For the purposes of this study we will consider only ROUGE-1 therefore a good ROUGE score is above 0.5 (Walker, n.d.).

ROUGE Metric	Excellent	Good	Moderate
ROUGE-1	0.5+	>0.5	0.4-0.5
ROUGE-2	-	>0.4	0.2-0.4
ROUGE-L	-	~0.4	0.3-0.4

Figure 8 (Walker, n.d.).

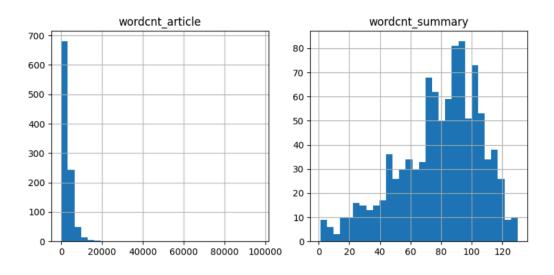


Figure 9 Figure 10

	wordcnt_article	wordcnt_summary
count	1000.000000	1000.000000
mean	2921.205000	79.036000
std	4677.075085	27.004703
min	8.000000	1.000000
25%	792.750000	64.000000
50%	2025.000000	84.000000
75%	3885.250000	99.000000
max	96769.000000	130.000000

Figure 11

After running the program we see the outputs of Fig 9,10 and 11. In Fig 9 and Fig 10 we see that the number of articles is over 800. This correlates to Fig 11 as the number of articles being displayed are 1000.

Demonstration of the wordcnt_summary for the whole dataset shows a mean of 79 words in comparison to the original articles word count of 2921. This is being run with 1000 files from the ToSDR dataset due to CPU limitations as it is being run on kaggle. This results in an average percentage reduction of 97.30% for abstractive summarization. Which is impressive compared to the average percentage reduction of 76.9% (Fig 17)when conducting extractive summarization on the whole dataset, however, there could be bias involved in this as I was able to run the full dataset which consisted of 9496 files when conducting extractive summarization. This was possible as extractive summarization is faster than abstractive summarization.

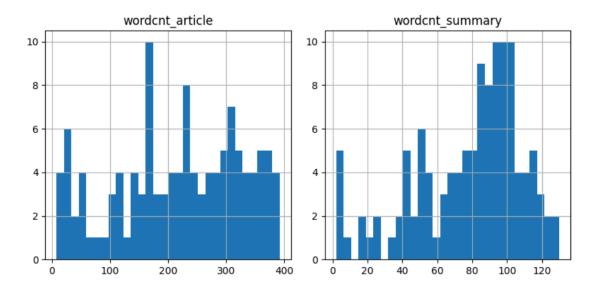


Figure 12 Figure 13

	wordcnt_article	wordcnt_summary
count	120.000000	120.000000
mean	217.750000	78.525000
std	108.531934	30.743986
min	8.000000	2.000000
25%	139.750000	56.750000
50%	226.000000	87.000000
75%	310.000000	100.250000
max	392.000000	130.000000

Figure 14

After looking at the word count distribution the dataset was reduced to those being less than 400 to stay within the models allowed max_length (Erturan, 2023). In Fig 14 We see that the number of articles is now 120. The reduction percentage has also dropped to 63.94%, which shows that the articles are being summarized less. This could be inaccurate as our dataset is too small now and shorter articles could be harder to summarize.

After running the aforementioned steps we are now able to demonstrate the ROUGE 1 precision score and the actual and predicted summary. In Table 1 based on the ROUGE score of 0.8 for our first sample we see that the ROUGE score is classified as Excellent (see Fig 8 for ROUGE classifications). However, the summary and predicted summary produced by BART are nonsensical. The predicted summary is the most difficult to comprehend as it was unable to filter through symbols. This result is also seen in other articles with different data sets such as Erturan's (2023) and Zhao, M. (2018).

1 Sample Output: Table 1 (See Appendix B for the full output)

	Article	Actual Summary	Predicted Summary	ROUGE- 1 Precisio n
0	summarize:. DMCA Please send DMCA takedown notices to our designated agent dmca@coomer.party, with the subject "DMCA (coomer.party)." Coomer reserves the right to evaluate the validity of notices recieved, and will ignore and discard notices with invalid, incorrect, or omitted informationA valid notice must includean electronic or physical signature of the person authorized to act on behalf of the owner of the copyright or other intellectual property interest.contact information of the copyright owner or an authorized agent (full name, mailing address, phone number, email).identification/location of the infringing material.identification/location of the original material claimed to have	a valid notice must include the following: d I DMCA Please send DMCA takedown notices to our designated agent dmca@coomer.party, with the subject "DMCA (coomer).party)." Coomer reserves the right to evaluate the validity of notices recieved, and will ignore and discard notices with invalid, incorrect, or omitted information.A valid noticemust include.a statement that the copyright owner or an authorized agent (full name, mailing address, phone number, email)hasan electronic or physical signature of the person authorized to act on behalf of the owner of the copyright or other intellectual property interest in the infringing materiala statementthat the person sending the notice has a good faith	summarize:. (""" "" Please send DMCA takedown notices to our designated agent dmca@coomer.party, with the subject "DMCA (coomer".party)." Coomer reserves the right to evaluate the validity of notices recieved, and will ignore and discard notices with invalid, incorrect, or omitted informationA valid notice must includean electronic or physical signature of the person authorized to act on behalf of the owner of the copyright or other intellectual property interest.a statement that the copyright owner or an authorized agent (full name, mailing address, phone number, email).identification/locati on of the infringing material.	0.80000

Discussion

An unintentional bias occurred as the program was run on Kaggle I had limitations of

working with only 20GB of data. When conducting the abstractive summarization the number of files that could be run had to be limited to 1000 due to my CPU and data constraints. There are also limitations regarding my scope of knowledge. As this was my first NLP program there are areas where the code could be more concise and clean. My chosen data set also posed a limitation, as I needed to create summaries by myself, my training data may be inaccurate. For future studies, this program could be run in an environment with a larger CPU and with larger storage. A dataset with training data could also be used to improve the output of the summarizations for abstractive methods.

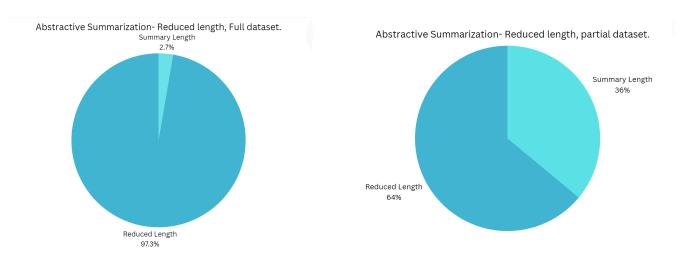


Figure 15

Extractive Summarization- Reduced length, Full dataset.

Summary Length
23.1%

Reduced Length
76.9%

Figure 17

The above pie chart showcases the percentage with which the model was able to summarize the files in the data set. As mentioned above, the extractive method Fig 17 ran all

9496 files in the dataset and had an average reduced length of 76.9%. While Fig 15 ran 1000 files and had an average reduced length of 97.3%. In comparison to this Fig 16 ran 120 files and had a reduced length of 64%. This demonstrates that Abstractive summarization can summarize the text more than extractive summarization when given a large dataset.

However, as the output for Abstractive summarization shows (see Appendix B). The Actual Summary and Predicted summary are nonsensical when compared to the summary that we received from extractive summarization (see Appendix A). While the reduction percentage of extractive summarization is quite low compared to abstractive summarization. It is higher in accuracy. In addition to this abstractive summarization took 4 hours and 25 minutes to run 1000 files instead of compared to the average run time of 48 minutes to run the extractive summarization for the whole dataset of 9496 files. Each file takes 0.3 seconds to run in extractive summarization. This was due to my need to create the summarization files for training myself as my data set did not include this information.

The abstractive method is also less user-friendly as it requires more theoretical knowledge and more time to run the program from different Kaggle accounts. As each time it is run on a different account, the summarizations need to be generated by BART.

The extractive method also provides additional data such as the top 20 words by frequency (Fig 3), word connections (Fig 4), keywords by PageRank Score (Fig 6), Distribution of Sentence Scores (Fig 7) and Distribution of Privacy Concerns Sentence Lengths (Fig 8). While this data is not crucial it could provide the user with additional information that they can use to make judgments on whether they want to agree to the website's TOS. The visual guides provided by these graphs could entice users to read the summary of the TOS.

Extractive summarization's outputs are also more applicable to the goal that we would like to achieve as a program. To assist users who find it difficult to read lengthy TOS documentation. The extractive summarization method allows for each website TOS file to

display the summarizations while in abstractive summarization data is processed and output in batches.

To summarize, extractive summarization is more applicable for our purpose as it's summarization is coherent, runs faster, is more user-friendly and provides additional visual data that could be useful for the user.

Conclusion

In conclusion, this research has highlighted the significant threat that AI data collection poses to privacy. Multiple sources have supported the notion that consumers express concerns regarding their privacy when using AI, due to unintentional data collection and potential misuse of sensitive information. The operant nature of AI can lead to the inference of data that users may not even know about themselves, raising further privacy concerns. While anonymizing data may help mitigate these risks, users should still be informed of potential risks and willingly choose to share their data.

To address these challenges politicians, programmers, and big corporations should come together and recognize the importance of maintaining privacy. Protecting user privacy builds trust and reduces the likelihood of future legal issues for companies. In addition, implementing new tools and methods to determine whether customers are willing to forego their privacy for certain benefits and services is crucial. The current method of relying on lengthy, complicated terms and services is ineffective and time consuming. Transparency and informed consent are key to addressing privacy concerns, and AI can automate this process, providing users with greater visibility into data collection and usage.

This study aimed to provide a solution by creating open-source software that uses abstractive and extractive text summarization methods to provide users with a summary of lengthy TOS

documentation to encourage users to read the critical parts of TOS of websites before agreeing to them. Extractive methodology was found to be more applicable to this goal as it provided coherent summaries, ran faster, is more user friendly and provided additional visual aids for users.

Ultimately, balancing utilizing data for marketing purposes and respecting consumer privacy is crucial. By prioritizing privacy and implementing transparent practices, stakeholders can ensure that AI's data collection is conducted ethically and responsibly, fostering trust between users and AI systems.

References

- Antón, A. I., Earp, J. B., & Young, J. D. (2010). <u>How internet users' privacy concerns have</u> evolved since 2002. *IEEE Security & Privacy*, 8(1), 21-27.
- Ben-Reuven, E. (2017, March 19). How_to_make_a_text_summarizer/README.md at

 master · llSourcell/How_to_make_a_text_summarizer. GitHub.

 https://github.com/llSourcell/How_to_make_a_text_summarizer/blob/master/READ

 ME.md
- Berton, R., Kolasinska, A., Gaggi, O., Palazzi, C. E., & Quadrio, G. (2020, September). <u>A chrome extension to help people with dyslexia</u>. In *Proceedings of the International Conference on Advanced Visual Interfaces* (pp. 1-5).
- Bleier, A., Goldfarb, A., & Doldfarb, A., & Doldfarb, A., & Doldfarb, C. (2019). Consumer privacy and the future of data-based innovation and Marketing. International Journal of Research in Marketing, 37(3), 466–480. https://doi.org/10.1016/j.ijresmar.2020.03.006
- Chintalapati, S., & Pandey, S. K. (2022). <u>Artificial intelligence in marketing: A systematic</u>
 <u>literature review.</u> *International Journal of Market Research*, 64(1), 38-68.
- Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). <u>How artificial intelligence will change the future of marketing</u>. *Journal of the Academy of Marketing Science*, 48, 24-42.
- Erturan, A. (2023, September 26). *Text_Summarization-bart-base*. Kaggle. https://www.kaggle.com/code/alkanerturan/text-summarization-bart-base

- Foxman, E. R., & Kilcoyne, P. (1993). <u>Information technology, marketing practice, and consumer privacy: Ethical issues</u>. *Journal of Public Policy & Marketing*, *12*(1), 106-119.
- Grewal, D., Hulland, J., Kopalle, P. K., & Karahanna, E. (2020). *The future of technology and marketing: A multidisciplinary perspective. Journal of the Academy of Marketing Science*, 48, 1-8.
- Grewal, D., Guha, A., Satornino, C. B., & Schweiger, E. B. (2021). <u>Artificial intelligence:</u>

 The light and the darkness. *Journal of Business Research*, 136, 229-236.
- Gupta, S. (2023, March 14). *TOSDR terms of Service Corpus*. Kaggle.

 https://www.kaggle.com/datasets/sonu1607/tosdr-terms-of-service-corpus/data
- Hettiaratchy, E. (2024a, July 22). *Abstractive Summarization with BART ToSDR*. Kaggle. https://www.kaggle.com/code/ericahh/abstractive-summarization-with-bart-tosdr
- Hettiaratchy, E. (2024b, July 22). *Extractive Summarization*. Kaggle. https://www.kaggle.com/code/ericahh/extractive-summarization
- Hsu, C. L., & Lin, J. C. C. (2016). <u>An empirical examination of consumer adoption of Internet of Things services: Network externalities and concern for information privacy perspectives</u>. *Computers in human behavior*, 62, 516-527. Vimal
- Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, 49, 30-50.
- Kim, D., Park, K., Park, Y., & Ahn, J. H. (2019). Willingness to provide personal

- Ethical and Privacy concerns for AI users in Marketing information: Perspective of privacy calculus in IoT services. Computers in Human Behavior, 92, 273-281.
- Kim, M. S., & Kim, S. (2018). <u>Factors influencing willingness to provide personal</u>
 <u>information for personalized recommendations</u>. *Computers in Human Behavior*,
 88, 143-152.
- Kumar, V., Rajan, B., Venkatesan, R., & Lecinski, J. (2019). <u>Understanding the role of artificial intelligence in personalized engagement marketing</u>. <u>California Management Review</u>, 61(4), 135-155.
- Learning to read and interpret network graph data visualizations. Nodus Labs: Ecological

 Thinking through Network Analysis. (2022, June 4).

 https://noduslabs.com/cases/learn-read-interpret-network-graphs-data-visualization/.
- Magdy, M. (2023, August 21). *Text-summarizer-bart-rouge-pytorch*. Kaggle.

 https://www.kaggle.com/code/mohamedmagdy191/text-summarizer-bart-rouge-pytorch/
 https://www.kaggle.com/code/mohamedmagdy191/text-summarizer-bart-rouge-pytorch/
- Mogaji, E., Soetan, T. O., & Kieu, T. A. (2020). <u>The implications of artificial intelligence on the digital marketing of financial services to vulnerable customers.</u> *Australasian Marketing Journal*, j-ausmj.
- Pitardi, V., & Marriott, H. R. (2021). <u>Alexa, she's not human but... Unveiling the drivers of consumers' trust in voice-based artificial intelligence</u>. *Psychology & Marketing*, 38(4), 626-642.

- Ethical and Privacy concerns for Al users in Marketing
- Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing*, 85(1), 131-151.
- Rao, A. (2024, February 5). *Text summarization in NLP*. GeeksforGeeks. https://www.geeksforgeeks.org/text-summarization-in-nlp/
- Roy, H., Borchardt, J.-C., Jong, M. de, McGowan, I., Stout, J., Azmayesh, S., O'Leary, M., Tunru, V., Talib, C., Mullen, E., Back, J. R., Cortés, G. M., Pedersen, D., Rubio, I., Barnes, A. J., Hering, E., Radkoff, E., & Beneke, M. (2012, June 1). *Terms of service; didn't read*. Frontpage. https://tosdr.org/
- Swain, S. (2023, June 8). *Bart model for text auto completion in NLP*. GeeksforGeeks. https://www.geeksforgeeks.org/bart-model-for-text-auto-completion-in-nlp/
- Varangouli, E. (2023, June 7). *Everything you need to know about google pagerank*. Semrush Blog. https://www.semrush.com/blog/pagerank/
- 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, 106763.
- Vlačić, B., Corbo, L., e Silva, S. C., & Dabić, M. (2021). <u>The evolving role of artificial intelligence in marketing: A review and research agenda</u>. *Journal of Business Research*, *128*, 187-203.
- Walker, S. M. (n.d.). What is the Rouge Score (recall-oriented understudy for GISTING evaluation)?. Klu. https://klu.ai/glossary/rouge-score

- Wang, E. S. T., & Lin, R. L. (2017). <u>Perceived quality factors of location-based apps on trust,</u>

 <u>perceived privacy risk, and continuous usage intention.</u> *Behaviour & Information Technology*, 36(1), 2-10.
- Zhang, X., Liu, S., Chen, X., Wang, L., Gao, B., & Zhu, Q. (2018). <u>Health information</u>

 <u>privacy concerns, antecedents, and information disclosure intention in online health</u>

 <u>communities</u>. *Information & Management*, 55(4), 482-493.
- Zhao, M. (2018, June 8). Extractive vs. Abstractive Text Summarization Methods: An Analysis. github.
- Zhou, W., Jia, Y., Peng, A., Zhang, Y., & Liu, P. (2018). The effect of iot new features on security and privacy: New threats, existing solutions, and challenges yet to be solved. *IEEE Internet of things Journal*, 6(2), 1606-1616.

https://github.com/mzhao98/text-summarization/blob/master/README.md

Ethical and Privacy concerns for Al users in Marketing Appendices

Appendix A

Example Output: Extractive summarization

Enter the file path (or 'exit' to quit): /kaggle/input/tosdr-terms-of-service-corpus/text/Juicer CookiePolicy.txt

Summary of the document:

It is recommended that you leave on all cookies if you are not sure whether you need them or not in case they are used to provide a service that you use. How to disable CookiesIf you do not want information collected through the use of cookies when using your Juicer account or our website, there is a simple procedure in most browsers that allows you to automatically decline cookies or gives you the choice of declining or accepting the transfer to your computer of a particular cookie (or cookies) from a particular site. We may also use cookies to see how you interact with our offers, and we may use cookies or other files to understand your use of our websites. How does Juicer. io use Cookies? We use cookies for a variety of reasons detailed below. For example, we may not be able to recognize your computer, and you may need to log in every time you visit the applicable Digital Services. The Cookies We SetAccount related Cookies If you register an account on Juicer.io we will use cookies to manage the signup process and general administration. Those cookies are deleted when you log out to ensure that other users of your machine cannot access your private data. Third Party CookiesIn some special cases we also use cookies provided by trusted third parties. These cookies will be deleted when you log out.Login related cookiesTo prevent you from having to log in every single time you visit Juicer.io, we use cookies to remember your login state.

Privacy Concerns identified in the document:

Cookie Policy for Juicer.ioWhat Are Cookies? Cookies are pieces of information stored directly on the computer that you are using. Cookies allow us to collect information such as browser type, time spent on the Digital Services, pages visited, language preferences, and other web traffic data. We, and our service providers, use the information for security purposes, to facilitate online navigation, to display information more effectively, to personalize your experience while using the Digital Services, and to otherwise analyze user activity. Cookies further allow us to select which of our offers are most likely to appeal to you and to display them to you. We may also use cookies to see how you interact with our offers, and we may use cookies or other files to understand your

use of our websites. How does Juicer. io use Cookies? We use cookies for a variety of reasons detailed below. Unfortunately in most cases there are no industry standard options for disabling cookies without completely disabling the functionality and features they add to this site. It is recommended that you leave on all cookies if you are not sure whether you need them or not in case they are used to provide a service that you use. How to disable CookiesIf you do not want information collected through the use of cookies when using your Juicer account or our website, there is a simple procedure in most browsers that allows you to automatically decline cookies or gives you the choice of declining or accepting the transfer to your computer of a particular cookie (or cookies) from a particular site. You may also wish to refer to

http://www.allaboutcookies.org/manage-cookies/index.html. However, if you do not accept these cookies, you may experience some inconvenience in your use of the Digital Services. For example, we may not be able to recognize your computer, and you may need to log in every time you visit the applicable Digital Services. The Cookies We SetAccount related CookiesIf you register an account on Juicer.io we will use cookies to manage the signup process and general administration. These cookies will be deleted when you log out.Login related cookies To prevent you from having to log in every single time you visit Juicer.io, we use cookies to remember your login state. Those cookies are deleted when you log out to ensure that other users of your machine cannot access your private data. Third Party CookiesIn some special cases we also use cookies provided by trusted third parties. The following section details which third party cookies you might encounter through this site. The payment processor Stripe is used to process your payments and cookies are essential to ensure that your order is remembered between pages so that Stripe and Juicer.io can process it properly. These cookies may track things such as how long you spend on the site and the pages that you visit so we can continue to produce engaging content. Zendesk utilizes cookies to keep track of ongoing and past conversations. When we are still testing new features these cookies may be used to ensure that you receive a consistent experience whilst on the site whilst ensuring we understand which optimisations our users appreciate the most. Facebook, Instagram and Twitter, will set cookies through our site which may be used to enhance your profile on their site or contribute to the data they hold for various purposes outlined in their respective privacy policies. More InformationHopefully that has clarified things for you and as was previously mentioned if there is something that you aren't sure whether you need or not it's usually safer to leave cookies enabled in case it does interact with one of the features you use on our site.

Original length of the document: 4715 characters Summarized length of the document: 1453 characters

Reduction in length: 69.18%

Appendix B

Example Output: Abstractive summarization

	Article	Actual Summary	Predicted Summary	ROUGE-1 Precision
0	summarize:. DMCA Please send DMCA takedown notices to our designated agent dmca@coomer.par ty, with the subject "DMCA (coomer.party)." Coomer reserves the right to evaluate the validity of notices recieved, and will ignore and discard notices with invalid, incorrect, or omitted informationA valid notice must includean electronic or physical signature of the person authorized to act on behalf of the owner of the copyright or other intellectual property interest.contact information of the copyright owner or an authorized agent (full name, mailing address, phone number,	a valid notice must include the following: d I DMCA Please send DMCA takedown notices to our designated agent dmca@coomer.par ty, with the subject "DMCA (coomer).party)." Coomer reserves the right to evaluate the validity of notices recieved, and will ignore and discard notices with invalid, incorrect, or omitted information.A valid noticemust include.a statement that the copyright owner or an authorized agent (full name, mailing address, phone number, email)hasan electronic or physical signature of the person authorized to act on behalf of the owner of the copyright or other	summarize:. (""" "" Please send DMCA takedown notices to our designated agent dmca@coomer.par ty, with the subject "DMCA (coomer".party)." Coomer reserves the right to evaluate the validity of notices recieved, and will ignore and discard notices with invalid, incorrect, or omitted informationA valid notice must includean electronic or physical signature of the person authorized to act on behalf of the owner of the copyright or other intellectual property interest.a statement that the copyright owner or an authorized agent (full name, mailing address, phone number, email).identification /location of the	0.800000

Ethical and Privacy	Ethical and Privacy concerns for AI users in Marketing				
	email).identification /location of the infringing material.identificati on/location of the original material claimed to have been infringed.a statement that the person sending the notice has a good faith belief that use of the material in the manner complained of is not authorized by the copyright owner, its agent, or the law.a statement that the information in the notice is accurate, and under penalty of perjury, that the person sending the notice is authorized to act on behalf of the copyright owner. Notices may be shared with 3rd parties for transparency purposesWe usually respond within 1-3 business days	intellectual property interest in the infringing materiala statementthat the person sending the notice has a good faith	infringing material.		
1	summarize:. Privacy Policy. No cookiesNo trackingNo collecting dataClick here to return to the home page	Privacy Policy Â No cookies on your device.No tracking of your activity.No collecting data.No sharing of personal information.Click here to return to the home page.	""e'e>" on Mau Mau Mau" he always, he's calm as he would be if he always would be so strong, heft he heh Mau Maucrested by "1" or "2" "E'ft> on "Oft" tied by "" Mauc> if he would always would> >> > on >	0.035714	

Ethical and Privac	y concerns for AI u	sers in Marketing		33
			Mau> he he> 0> for " > " on > tied" by "", "3rd "ribed by >" "P	
2	summarize:. Find more solutions Digital Services and Device Support â ° Amazon Device and Digital Services Terms, Warranties, and Noticesâ ° Kindle Special Offer Text Notifications Terms & Samp Conditions The short code 262966 is used by Amazon.com to send automated text notifications (U.Sonly) If you agree to receive text notifications from Amazon, you will receive text messages when there is an upcoming limited-time, limited-quantity deal exclusively available on your Kindle Fire with Special Offers Messages will be sent between 8am - 6pm PT Frequency varies Message and data rates may apply You may opt-out at any time by sending STOP to 262966 The following carriers support Amazon	- Find more solutions Â Amazon Device, Digital Services and Device Support â ° Amazon Device and Digital Services Terms, Warranties, and Noticesâ f° Kindle Special Offer Text Notifications Terms & Warranties Please see our Privacy Notice for how we use your information. Â	summarize:. Find more solutions Digital Services and Device Support â ° Amazon Device and Digital Services Terms, Warranties, and Noticesâ¡° Kindle Special Offer Text Notifications Terms & Services\$ampCo nditions The short code 262966 is used by Amazon.com to send automated text notifications (U.Sonly)If you agree to receive text notifications from Amazon, you will receive text messages when there is an upcoming limited-time, limited-quantity deal exclusively available on your Kindle Fire with Special OffersMessages will be sent between 8am - 6pm PTFrequency variesMessage and data rates may applyYou may opt-out at	0.849462

Ethical and Privacy	<u>concerns for Al u</u>	sers in Marketing	
	text notifications:.		
	Alaska	Messages will	
	Communications	be sent between	
	Systems (ACS),		
	Alltel-Allied, AT&T,		
	Bluegrass Cellular,		
	Boost Mobile,		
	Boost-CDMA, C		
	Spire Wireless		
	(aka Cellular		
	South), CellCom,		
	Chariton Valley		
	Cellular, Chat		
	Mobility, Cincinnati		
	Bell, Cricket (Leap		
	Wireless), Cross		
	Wireless, DTC		
	Wireless, Duet IP		
	(MaxWireless/Wir		
	eless		
	CommVenture),		
	East Kentucky		
	Network		
	(Appalachian		
	Wireless), ECIT,		
	Element Mobile,		
	Epic Touch, GCI		
	Communications,		
	Golden State		
	Cellular, Illinois		
	Valley Cellular,		
	Immix, Inland		
	Cellular, IWireless,		
	Metro PCS		
	(including		
	subsidiary		
	PTA-FLA),		
	MetroPCS/T-Mobil		
	e US, Mobi PCS,		
	MobileNation/SI		
	Wireless, Mosaic		
	Telecom, MTA		
	Wireless/Matanusk		
	a Kenai, MTPCS		
	Cellular One		
	(Cellone Nation),		
	Nex Tech		
	Communications, Northwest Missouri		
	Cellular, nTelos,		

Ethical and Privacy	concerns for AI us	sers in Marketing	
	Panhandle		
	Wireless, Peoples		
	Wireless, Pine		
	Cellular, Pioneer		
	Cellular, Plateau		
	Wireless, Revol		
	Wireless, Rural		
	Independent		
	Network Alliance		
	(RINA),		
	SouthernLINC,		
	Sprint, SRT		
	Communications,		
	Thumb Cellular,		
	T-Mobile, TracFone		
	(AT&T), Union		
	Telephone, United		
	States Cellular		
	Corp, United		
	Wireless, Verizon,		
	Viaero Wireless,		
	Virgin Mobile, West		
	Central		
	Wireless.Please		
	note that you will		
	not be able to		
	receive text		
	messages on VoIP		
	numbers like		
	Google Voice or		
	Skype What		
	Will Amazon Do		
	with My Phone		
	Number?.Please		
	see our Privacy		
	Notice for how we		
	use your		
	information Was		
	this information		
	helpful? Yes No		
	Thank you for your		
	feedback Please		
	select what best		
	describes the		
	information: This		
	information is		
	confusing or		
	wrong. This isn		

summarize:. ParticipationintheO ffering(asdefinedbe low)involvesahighd egreeofrisk. Yousho uld,carefullyreviewt heconfidentialinfor mationprovidedtoy ouinconnectionher ewith,,togetherwith WIFI-Tokens,To alloftheotherinform *kenPurchaseTer* summarize:. �� ationcontainedinth ms(asdefinedbelo **Participationinthe** eseTerms(asdefine Offering(asdefine w),,beforemaking dbelow),,beforema dbelow)involvesa apurchasedecisio kingapurchasedeci highdegreeofrisk. sion., Youshould,careful lyreviewtheconfid Wifl-Token, Token WIFI-Tokens, To entialinformation PurchaseTerms,1. providedtoyouinc kenPurchaseTerPurchasing,theTo onnectionherewit ms,1..General,W kensherein,Sectio h,,togetherwithall eConnectUAG,cl n3belowaswellasi oftheotherinforma *oPSTConsulting* ntheLitepaperby,S tioncontainedinth GmbH,Baarerstr ection4below,care eseTerms(3 0.320000 fullyreviewthecon asdefinedbelow),, asse10,6300Zug fidentialinformati beforemakingapu ,Switzerland(the" onprovidedtoyoui rchasedecision.,\$ Company"),andl nconnectionhere WIFI-Tokens, Toke oritsaffiliates,am with,,togetherwith nPurchaseTerms, ongotherthings,d alloftheotherinfor 1..G eneral, We C evelopsoftwarere mationcontainedi onnectUA G,c/oPST C ntheseTerms.Inth latingtoblockchai onsulting G atcase, aftermakin nordistributedled gappurchaseddec mbH,Baarerstrass ger,technologyor ision.,,before e10,6300Zug,Swit a similar technolozerland(the" C makingapurchasi gy,includingaplat ngdecision.\$Purc ompany"),and/orit saffili *formprovidingWi* hasedCaps,Purch asers,TokenPurch Fiandotherservic ases,Tokens,Secti esrelatingtodata, on5belowasselfas (the**"**Project wellAswellas **â€**).Uponthetok engeneratingeve nt,atotalof1' 000'000â€ ™000(onebillion)tokenswillbe,cre ated(the**"**Tok

Ethical and Privac	y concerns for Al u	sers in Marketing		59
Limear and Titrae		Sers in warketing		
	enGeneratingEve			
	nt †). Thefollow			
	ingterms(the"Ter			
	ms ") relatetothep			
	ublicoffering,(the			
	" Offering â€			
)ofupto115,000,			
	000			
	000			
	WIFI-Tokens(theâ			
	⊛Tokenâ)witht			
	hefunctionalitiesa			
	sdefinedin,Sectio			
	n3belowaswellasi			
	ntheLitepaperby			
	We C onnectUA G			
	(asamendedfromt			
	imetotime)(the,â			
	∞Litepaperâ).Th eOfferingwilltake			
	placefrom1Octob			
	er2022until14Oct			
	ober2022(the"Off			
	erPeriod").,These			
	Termsdonotconsti			
	tuteaprospectuso			
	fanysort,norasoli			
	citationorrecomm			
	endationfor,invest			
	ment,noranofferin			
	gofsharesorequit			
	y,nordotheypertai			
	ntoanofferingofse			
	curitiesinanyform			
	,ormanner.,2.Sco			
	peofTerms,These			
	Termsgovernyour			
	purchaseofToken			
	sfromthe C			
	ompanyduringthe			
	OfferPeriod.Bypu			
	rchasing,theToke			
	nsherein,you(theâ			
	@ P			
4	summarize:.We	We sometimes	summarize:.We	0.633663
	sometimes place	place small data	sometimes place	
	Jointaines place	piaco sinan data	Jointaines place	

small data files on your computer..These are known as cookies..Most websites do this..They improve things by:..Rememberin g settings, so you donâ ™t have to keep re-entering them whenever you visit a new page remembering information youâ ™ve given (eg your postcode) so you donâ ™t need to keep entering it.. Measuring how you use the website so we can make sure it meets your needs...Our cookies arenâ ™t used to identify you personally..theyâ ™re just here to make the site work better for you..Indeed, you can manage and/or delete these small files as you wish..To learn more about cookies and how to manage them, visit AboutCookies.org ..Or please read on to find out more about how and where we use cookies..How we

files on your computer.These are known as cookies.We use cookies in several places.Most websites do this.They improve things by:Protecting user privacy.Protecting your privacy.Remembe ring settings, so you donâ ™t have to keep re-entering them whenever you visit a new page remembering information about the site.Protects your privacy by remembering details about the page you visit, such as your name or address.Passwor dProtecting information about your account.Passwor dRemembering information youâ ™ve given (eg your postcode) so you can donâ t need to keep entering it.Passwordprote us.PasswordMea suring how

small data files on your computer..These are known as cookies..Most websites do this..They improve things by:..Rememberin g settings, so you donâ ™t have to keep re-entering them whenever you visit a new page, you don't need to keep entering it..Measuring how you use the website so we can make sure it meets your needs...Our cookies arenâ œ™re just here to make the site work better for you..These, you can manage and/or delete these small files as you wish..To learn more about cookies and how to manage them, visit AboutCookies.org ..Or please read

ai and Privacy	y concerns for AI us	sers in Marketing	
	use cookies.We		
	uses cookies in		
	several places		
	â " weâ ™ve		
	listed each of		
	them below with		
	more details		
	about why we use		
	them and how		
	long they will		
	lastln most		
	cases, we will let		
	you know before		
	we store a cookie		
	on your computer		
	by providing a		
	messageMeasuri		
	ng website usage		
	(Google		
	Analytics).We use		
	Google Analytics		
	to collect		
	information about		
	how people use		
	this siteWe do		
	this to make sure		
	itâ ™s meeting its		
	usersâ ™ needs		
	and to		
	understand how		
	we could do it		
	betterGoogle		
	Analytics stores		
	information about		
	what pages you		
	visit, how long		
	you are on the		
	site, how you got		
	here and what		
	you click onWe		
	do not collect or		
	store your		
	personal		
	information		
	(e.gyour name		
	or address) so this information		
	cannot be used to		
	identify who you		
	areWe do not		
	areve ao not		

Ethical and Privacy concerns for AI users in Marketing					
	allow Google to use or share our analytics data.				

© 2024

Erika M. Hettiaratchy ALL RIGHTS RESERVED