INSTITUTO
UNIVERSITÁRIO
DE LISBOA

TRUST FORMATION IN GENERATIVE AI: THE ROLE OF MEDIA RICHNESS AND VALUE SYSTEMS

Alina Flekler

Masters of Science in Marketing

Supervisor:
Prof. João Guerreiro, Assistant Professor, ISCTE-IUL Business School, Department of Marketing, Operation and General Management.

September, 2024

BUSINESS
SCHOOL

TRUST FORMATION IN GENERATIVE AI: THE ROLE OF MEDIA RICHNESS AND VALUE SYSTEMS

Alina Flekler

Masters of Science in Marketing

Supervisor:
Prof. João Guerreiro, Assistant Professor, ISCTE-IUL Business School, Department of Marketing, Operation and General Management.

September, 2024

Resumo

Esta tese investiga o impacto da riqueza de meios de comunicação na confiança dos usuários finais em IA generativa, com foco no papel mediador dos Valores. À medida que a IA generativa se expande por diferentes setores, compreender os fatores que moldam a confiança pública é essencial para sua aceitação e implementação eficaz. Baseado na Teoria da Riqueza de Mídia, o estudo hipotetiza que uma maior Riqueza de Mídia—caracterizada por maior conteúdo e qualidade—melhora a confiança em marcas e sistemas de IA generativa. O estudo também examina como os Valores mediam essa relação, proporcionando uma visão detalhada da dinâmica da confiança. O estudo empírico, baseado nas respostas de 300 consumidores e utilizando a modelo de equações estruturais por mínimos quadrados (PLS-SEM), confirma um efeito positivo significativo da riqueza de mídia nos níveis de confiança. Foi constatado que os Valores desempenham um papel mediador em como a Riqueza de Mídia influencia a Confiança em IA generativa. Esses achados ressaltam a importância de projetar comunicações de IA que sejam ricas em conteúdo e alinhadas com os Valores do público, ao mesmo tempo em que gerenciam cuidadosamente a quantidade de informações técnicas fornecidas. Para desenvolvedores e profissionais de marketing de IA, os achados enfatizam a necessidade de integrar princípios de design orientados por valores nos sistemas de IA e nas estratégias de comunicação.

Palavras-chave: IA generativa, riqueza de meios de comunicação, confiança em marcas, confiança em IA, mediação de valores, ansiedade tecnológica, percepções dos usuários finais, conteúdo informativo, comunicação orientada por valores

JEL Sistema de Classificação: M31, M37, D01

Abstract

This thesis investigates the impact of Media Richness on end-user trust in generative AI, focusing on the mediating role of Values. As generative AI expands across sectors, understanding the factors that shape public trust is essential for its acceptance and effective implementation. Based on Media Richness Theory, the study hypothesizes that higher Media Richness—characterized by greater content and quality—enhances trust in brands and generative AI systems. It also examines how Values mediate this relationship, providing a nuanced view of trust dynamics. The empirical study based on responses from 300 consumers and employing partial least squares structural equation modeling (PLS-SEM) confirms a significant positive effect of Media Richness on trust levels. foundVValues play a mediating role, particularly in how Media Richness influences Trust in generative AI. These findings highlight the importance of designing AI communications that are rich in content and aligned with audience Values, while carefully managing the amount of technical information provided. For AI developers and marketers, the findings emphasize the need to integrate value-driven design principles into AI systems and communication strategies.

Keywords: generative ai, Media Richness, Brand Trust, AI trust, values mediation, technology anxiety, end-user perceptions, information content, value-driven communication

JEL Classification System: M31, M37, D01, D12, D83

Acknowledgments

This dissertation is the result of countless hours of hard work, a rollercoaster of ups and downs, and a deep curiosity about a topic I'm truly passionate about. However, this journey would not have been possible without the support and contributions of several amazing people, to whom I owe my deepest gratitude.

First and foremost, I want to thank Professor João Guerreiro for his invaluable guidance throughout this research. His availability, dedication, patience, and insightful advice were crucial in shaping this dissertation and making it a truly valuable learning experience.

I also want to extend my heartfelt thanks to my company, Ella Lab, for not only providing financial support during this research but also for helping me bridge my interests in marketing and AI. A special thanks to Carina Rothkegel for her invaluable input, which was instrumental in grounding my thesis in scientific research while addressing a real-world business problem.

I'm really grateful to my family, friends and colleagues for their unwavering support, patience, and invaluable feedback. Their emotional support has meant the world to me. Thanks for being there through the ups and downs and for motivating me to keep going. Your support and belief in me have made this journey much more rewarding and manageable.

Finally, I want to extend my sincere thanks to everyone who participated in the questionnaire. Your contributions were vital to this research, and I truly appreciate your time and effort.

THE MEANING OF TRUST

"In the quest for understanding AI, trust serves as the bridge that transforms complexity into clarity." (T. Miller, 2019)

"Trust is the cornerstone of a resilient future, enabling progress, ensuring that advancements are embraced with confidence." (Putnam, 2000)

Table of Contents

List of Tables

List of Figure

Glossary of Acronyms

AI – Artificial Intelligence

BT – Brand Trust

GAI – Generative Artificial Intelligence

GT – Generative Artificial Intelligence Trust

IEP – Importance of Ethics Principles

MR - Media Richness

MRT - Media Richness Theory

1 Introduction

1.1 Relevance of the topic

The rise of generative artificial intelligence (GAI), described as a disruptive technology, has the potential to revolutionize how we live, work, and evolve as a society - generating excitement, raising questions, inducing anxiety (Berg et al., 2023a; Zhang & Gosline, 2023a). The technology is expected to significantly enhance human capabilities at a low cost (K. Liu & Tao, 2022; Schwab, 2017; Yi et al., 2006), infiltrate numerous industries, contribute $15.7 trillion to the global economy by 2030 (Murphy et al., 2021), increase global GDP by 7%, and potentially displace 300 million knowledge worker jobs (Goldman Sachs, 2023). Its application presents significant societal opportunities and challenges, introducing ethical complexities such as concerns about fairness, transparency, privacy, and human rights to (Ashok et al., 2022; El Atillah, 2023; Helbing et al., 2019; Kuziemski & Misuraca, 2020).

Although the exact impact of artificial intelligence (AI), the fundamental technology behind GAI, on the future is still uncertain, incidents such as a man taking his own life after an AI chatbot urged him to do so illustrate how AI can lead to fatal outcomes (El Atillah, 2023). If not properly managed, integrating AI into global society could pose catastrophic risks to humanity (Bostrom, 2014)**.** However, the development progresses regardless of individual approval or disapproval (Brynjolfsson & McAfee, 2014). The critical issue is not whether to engage with this disruptive process, but how to shape and guide it effectively based on commonly agreed human values (Brynjolfsson & McAfee, 2014; W. Li et al., 2023; Medaglia et al., 2023). The possible harmful effects of using AI systems have led to a lack of trust by users (Bach et al., 2024). Trust in AI and its providers is a critical factor in AI acceptance, a crucial predictor of both behavioral intentions and the use of AI technology (Choung et al., 2023c; Lancelot Miltgen et al., 2013; Thomas et al., 2013; Venkatesh et al., 2003; Zerilli et al., 2022). It enables users to feel confident that a device will fulfill its intended purpose, thereby affecting their willingness to adopt and utilize AI technology (Chang et al., 2017).

Trust perceptions are a key measure for understanding human value systems, and, consequently, for driving AI adoption rates, highlighting the importance of understanding the user perspective (Afroogh et al., 2024; Ahmed et al., 2023; Berente et al., 2021). Given the absence of governmental regulations at the time of this research, ethical guidelines serve as the primary principles for organizations, attempting to provide a framework for AI development and deployment to align with societal values (David et al., 2024; Hagendorff, 2020). However, these abstract principles require significant effort to apply practically, prompting calls to move from principles to practice (Bartoletti, 2020; Jobin et al., 2019; Mökander et al., 2021; Munn, 2023). Despite many existing frameworks, their effectiveness remains largely unknown (Stamboliev & Christiaens, 2024).

To adequately address society's informational needs, these frameworks must be tested in real-life scenarios using a user-centric approach (Afroogh et al., 2024; Choung et al., 2023; David et al., 2024), as the influence of trust on AI acceptance varies with context and the information available to users (Kelly et al., 2023). The key challenge is communicating technical topics in a trust-building manner that is understandable to non-experts (Johnson-Eilola & Selber, 2023; Liao & Sundar, 2022; Morley et al., 2020). It is essential to provide enough information to ensure safety without overwhelming users or causing distrust, similar to the negative effects of 'greenwashing'(Reinhardt, 2023; Wright, 2024). Achieving this balance is crucial for building trust without causing confusion or skepticism (Wacksman, 2021).

The rapid, unpredictable, and difficult-to-control development of GAI, along with its significant economic and societal potential and risks that will impact both professional and personal lives, makes understanding users' trust perceptions and creating robust trust strategies crucial for businesses, policymakers, and public institutions (Afroogh et al., 2024; Fügener et al., 2021). It is essential to understand: How can patients be encouraged to use AI health monitoring apps, employees to engage with AI performance review systems, parents to adopt AI educational tools, users to trust AI financial advisors, citizens to interact with AI chatbots for public services, and consumers to share personal data with AI-driven customer service platforms? Effective marketing strategies can address this issue and will be essential for communicating GAIs advantages and addressing potential concerns, thereby improving trust and adoption (Abillama et al., 2021; Berg et al., 2023a; Berglind et al., 2022; Murphy et al., 2021; Riparbelli, 2024; van Wynsberghe, 2021; Y. Zhang & Gosline, 2023a; Zhanga & Hußmann, 2021).

However, it is necessary to determine if the communication methods used to promote GAI adoption are effectively meeting their intended purpose (Daft & Lengel, 1986; Hollingshead & McGrath, 1993) to efficiently guide resource investment into human-centered AI (Kelly et al., 2023).

1.2 Problem statement

Despite growing recognition of trust as a pivotal element in user acceptance of AI systems (Devitt, 2018; Siau & Wang, 2018), there is a significant gap in understanding how organizations can effectively communicate their trustworthiness to meet end-users' needs. Existing research has predominantly concentrated on technological development and governance aspects of trust in AI (Liao & Sundar, 2022; Wafa & Muzammil Hussain, 2021), often overlooking the broader public discourse on trust and the necessity for a user-centric approach (Bach et al., 2024; Ziegler & Donkers, 2024). However, inconsistencies in how AI is defined and operationalized in research, coupled with difficulties individuals face in understanding AI, GAI, and related technologies, create challenges in comparing studies. For example, Liang and Lee (2017) Liang & Lee (2017)found confusion between autonomous robots and AI, which complicates deriving actionable insights for influencing AI acceptance (Kelly et al., 2023).

Additionally, traditional dimensions of trust, such as integrity, competence, and dependability, as discussed in strategic communication literature (Hon & Grunig, 1999; Paine, 2003; Shockley-Zalabak & Ellis, 2006), do not adequately address the complexities involved in building trust in AI. Trust in the technology itself is intertwined with trust in the organizations that provide it, and effective communication plays a critical role, especially in situations with a lot of uncertainty, such as the rapid development of generative AI (David et al., 2024; Luhmann, 1979; Mayer et al., 1995).

There is a notable deficiency in comprehensive research exploring the interplay between end-users' informational needs, their value systems, and trust perceptions in AI contexts (Mirbabaie et al., 2022). The integration of human value systems is crucial for fostering trust, with transparency, fairness, and moral responsibility being essential components (Adomako & Nguyen, 2023; Hadj, 2020; K. W. Miller et al., 2010; Prahl & Goh, 2021; Rahwan, 2018; Taylor

& Taylor, 2021; Xie et al., 2024). Previous studies have largely focused on isolated aspects of brand or AI trust, neglecting the integrative approach required to understand their combined impact (Botha & Pieterse, 2020; Clayton, 2023). Furthermore,while the concept of Media Richness has been extensively examined in traditional communication settings, its application to developing effective trust-building strategies for generative AI is still relatively unexplored (Ishii et al., 2019; Kock, 2005; Maity et al., 2018a).

Current literature highlights the importance of marketing research addressing end-user information needs and human-centered values to successfully manage the implementation of generative AI (Ahmad et al., 2023; Choung et al., 2023c; David et al., 2024; Grewal et al., 2017; Haleem et al., 2022; M.-H. Huang & Rust, 2018; Kelly et al., 2023; V. Kumar & Shah, 2004; Liao & Sundar, 2022; Mirbabaie et al., 2022).

To address these gaps, the approach will involve testing principles of trust and value by examining how Media Richness in generative AI communication influences trust, with a focus on the mediating role of personal values. Exploratory studies are necessary to observe real-world AI interactions and enhance the external validity of theoretical models beyond well-studied models like the Technology Acceptance Model (TAM) (Kelly et al., 2023; Ofosu-Ampong, 2024; Ahmad et al., 2023).

1.3 Research purpose

This study addresses gaps in understanding how Media Richness impacts Brand Trust and generative AI (GAI) trust, and how values mediate these relationships. By integrating theories from Media Richness, Brand Trust, and GAI trust, the research aims to provide a comprehensive view of their interactions. Utilizing the European Commission guidelines for trustworthy AI, the study will test a theoretical framework against user perceptions through the creation and evaluation of three distinct versions of a value-based company statement document for a fictional generative AI provider. Each version will feature varying levels of detail to assess how comprehensiveness and transparency affect users' trust perceptions.

The research seeks to advance both academic theory and practical marketing applications by examining how Media Richness can be leveraged to build and maintain trust in AI technologies. This will offer valuable insights for enhancing trust in AI systems and the

organizations behind them. The study aims to bridge the gap between theory and practical application, providing relevant insights for businesses, technology developers, policymakers, and those working at the intersection of generative AI and human interaction. A key objective is to ensure that AI development aligns with societal values, making the implementation process safer and more ethically sound.

Specifically, this study seeks to:

1. Examine the impact of Media Richness on Brand Trust.

2. Investigate the influence of Media Richness on GAI trust.

3. Explore the mediating role of values in the relationships between Media Richness, Brand Trust, and GAI trust.

By achieving these objectives, the research will provide valuable insights for marketing professionals and businesses, helping them develop trust building strategies for generative AI while still adhering to societal values and ethical standards. This alignment is crucial for fostering safer, more responsible AI implementations and promoting positive engagement with technology.

1.4 Research questions

The main question guiding this research is: *How much information do end-users need in order to build trust in AI and the organizations behind it, and what role do their value systems play in this?* To achieve the research purpose, the following research questions are proposed:

RQ1: How much information do end-users need to build trust in GAI and the organizations behind it?

RQ2: What values are important to people when it comes to trustworthy GAI design?

RQ3: What role do these values play in the communicative process of building trust in generative AI

1.5 Research outline

This thesis is organized into six main chapters, each designed to systematically address the research questions and objectives set forth. The structure is as follows:

(1) Introduction: The introductory chapter provides an overview of the research topic, emphasizing its relevance and outlining the specific problems being addressed. It also includes the research purpose, questions, and the overall structure of the thesis.

(2) Literature Review: This chapter delves into existing literature on generative AI, Media Richness, Brand Trust, and the role of values. It identifies gaps in current knowledge and establishes the theoretical foundation for the research.

(3) Methodology: This section outlines the research design, detailing the experimental setup, data collection methods, and analytical techniques employed to test the hypotheses. Results: The results chapter presents the findings from the data analysis, focusing on the relationships between Media Richness, trust, and values.

(4) Discussion: In this chapter, the results are interpreted within the context of existing literature, exploring their theoretical and practical implications. This chapter also addresses the limitations of the study and provides suggestions for future research.

(5) Conclusion: The final chapter summarizes the key findings, contributions to knowledge, and practical implications of the research. It concludes with recommendations for future research and practice.

2 Literature Review

2.1 AI: Disruptive Technology, Disrupted Worldviews

2.1.1 Understanding Generative AI

The term GAI refers to a subset of artificial intelligence (AI) that focuses on automatically creating new data and content that closely resembles human output (Mondal et al., 2023). AI essentially mimics human intelligence using computer systems (Berente et al., 2021). The roots of AI trace back to foundational frameworks established as early as 1950, aiming to create machines mirroring human faculties like perception, logic, and cognition (Berente et al., 2021; H. Wang et al., 2019). Over the past fifty years, technological advancements, coupled with vast datasets and improved algorithmic and computational capabilities, have propelled AI to become a significant force across various domains (Berente et al., 2021; Brynjolfsson et al., 2023; Taddeo & Floridi, 2018a). One definition of modern AI is ‘a machine-based system that can make predictions, recommendations, or decisions influencing real or virtual environments, given a set of human-defined objectives‘ (Choung et al., 2023a). According to Carvalho et al. (2019), AI is a collection of technologies that rely on processes like machine learning, natural language processing, and knowledge representation. Building upon Mirbabaie et al. (2022), this research views artificial intelligence not as a singular technology or a mere assortment of specific technologies. Instead, the definition of Berente et al. (2021) of AI as "the *frontier of computational advancements that references human intelligence in addressing ever more complex decision-making problems”* is adopted.

Generative AI is a specialized technology within the broader field of AI. Similar to AI, GAI is an umbrella term that refers to a class of algorithmic techniques (Strobel et al., 2024). ChatGPT, DALL-E, Midjourney and GitHub Copilot represent some of these systems (Chiu, 2024; Strobel et al., 2024). The essential distinction between GAI and traditional AI lies in the transition from predefined rule-based systems to GAI systems capable of autonomously generating novel content, ideas, and solutions (Kar et al., 2023) . GAI technologies use deep generative models to create new datasets from existing data (Boscardin et al., 2023).

In 2017, The Economist declared that data had overtaken oil as the world's most valuable resource (2017). GAI, being not only a data-driven but also a data-generating technology, is set to revolutionize human work and daily life alike (Y. Zhang & Gosline, 2023b). Academic and public discussions reflect a mixture of concern and excitement about this emerging technology (Botha & Pieterse, 2020; Clayton, 2023; Haupt & Marks, 2023; Khan, 2023; R. Li et al., 2023). However, the globe is keen to know how businesses, industries and societies will be impacted by GAI, such as ChatGPT-4 (Berg et al., 2023b).

Positioned as a disruptive technology capable of profoundly transforming our society, GAI has become a dominant topic in recent research and business,.(Ågerfalk, 2020; Fügener et al., 2021; Peres et al., 2023; Strobel et al., 2024). Leading the way in technological innovation, GAI is frequently seen as a key component of both the Fourth Industrial Revolution (Industry 4.0) and the impending Fifth Industrial Revolution (Industry 5.0) (Ahmed et al., 2023; French et al., 2021; Makridakis, 2017). The Fourth Industrial Revolution, or Industry 4.0, emphasizes the integration of digital technologies, such as AI and IoT, into manufacturing and other industries. The Fifth Industrial Revolution, or Industry 5.0, is expected to focus on the collaboration between humans and advanced technologies, aiming to enhance human capabilities and create more personalized and sustainable solutions.

This study specifically examines the perception of generative AI rather than general AI due to its distinctive ability to create new content and data, which affects user perceptions in unique ways. Generative AI, particularly through applications like ChatGPT, has gained widespread attention and become a prominent part of daily life, significantly influencing public perceptions of the technology (Mogavi et al., 2023). By focusing on generative AI, this study aims to provide more precise insights into public attitudes and industry impacts, thereby deepening the understanding of how different AI technologies are perceived and their broader societal effects (Ahmed et al., 2023; French et al., 2021; Makridakis, 2017). However, since generative AI is a new subset of AI and most of the existing literature uses AI as a broad term, insights are drawn from this broader body of research. Within the scope of this study, based on the work of Feuerriegel et al. (2024), GAI *refers to a type of artificial intelligence that possesses the ability to understand, learn, and apply knowledge across a wide range of tasks at a level comparable to human intelligence. Unlike previous narrow AI, which is designed for specific tasks, GAI aims to perform any intellectual task that a human can, exhibiting the ability to*

reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience.

The capacity to enable the creation and reimagination of content, as well as providing nearly human-like assistance holds significant consequences across various industries such as healthcare, entertainment, art, fashion, programming, sales and accounting (Bandi et al., 2023; Castelli & Manzoni, 2022; Mondal et al., 2023). AI is anticipated to define the next decade by significantly enhancing human capabilities at a low cost (K. Liu & Tao, 2022; Schwab, 2017; Yi et al., 2006). Projections suggest AI will revolutionize numerous industries, potentially contributing an estimated US $15.7 trillion to the global economy by 2030 (Murphy et al., 2021). According to industry reports, generative AI has the potential to increase global GDP by 7% and could lead to the displacement of 300 million jobs held by knowledge workers (Goldman Sachs, 2023). This technology promises to improve everyday life through advancements in healthcare, such as early detection (Becker, 2018), personalized customer service assistants (Murphy et al., 2021), tailored educational aids (Kashive et al., 2020), and automated transportation systems (Kaye et al., 2020), among other applications.

The current systems have advanced to the point where their outputs are difficult to distinguish from human creations (Mondal et al., 2023). AI for example already outperforms humans in terms of speed and accuracy across multiple areas (Jussupow et al., 2020). They excel in tasks such as analyzing CT images (Cheng et al., 2016), making judicial decisions (Highhouse, 2008) and predicting employee performance more effectively than traditional methods. Algorithms in the form of chatbots are starting to replace humans in customer service roles (Luo et al., 2019). Looking forward, machine learning capabilities are enabling algorithms to autonomously learn new games (Silver et al., 2017) and operate vehicles without human control (Bojarski et al., 2016) suggesting that AI autonomy may soon become a reality (Jussupow et al., 2020). As GAI stands to enhance society in fields like transportation (Xia et al., 2020), mental health care (Doraiswamy et al., 2019), and education (Modapothula et al., 2022), it is crucial to investigate the factors that promote its acceptance and adoption (Schmidt et al., 2021; Sohn & Kwon, 2020; Taddeo & Floridi, 2018b; Turner et al., 2010).

2.1.2 Navigating Change – Societal AI Phobia

Presently we find ourselves situated in a period wherein the global community is deeply concerned about the rapid spread of technologies and their potentially harmful effects (Ferreira et al., 2022). The integration of AI represents a transformative jump involving not just technological advancements but also cultural shifts, operational changes, and workforce dynamics (Agarwal, 2018; Ashok et al., 2016). Its application introduces ethical complexities, including concerns about fairness, transparency, privacy, and human rights (Ashok et al., 2022; Helbing et al., 2019; Kuziemski & Misuraca, 2020).

Despite the growing adoption of AI, there are increasing concerns about certain characteristics of these systems, such as their opacity (often referred to as "black box"), lack of interpretability, and inherent biases, along with the associated risks (Gulati et al., 2019; Rai, 2020). Distrust in AI, distinct from distrust in the organizations behind it, can stem from fear (Prahl & Goh, 2021, p. 6) and portrayals in science fiction (Devitt, 2018; Siau & Wang, 2018), particularly with innovations that are both radical and complex (Devitt, 2018). For example, films like *The Matrix* series show AI evolving beyond human control with detrimental effects on humans. Additionally, distrust is fueled by legitimate concerns about privacy, security, and accountability, particularly in the context of real-world issues such as system failures and AI biases (Prahl & Goh, 2021).The complex nature of AI complicates predictions about its behavior (Bathaee, 2018; Fainman, 2019), makes error tracing and decision backtracking more challenging (Bathaee, 2018), and hinders understanding the rationale behind its outputs (Fainman, 2019). These challenges are worsened by the inherent uncertainty in AI outputs (Zhanga & Hußmann, 2021). If AI systems are poorly designed or not properly tailored to users, they can lead to unfair and incorrect decision-making (Lakkaraju & Bastani, 2020). The real-world impact of failures in AI-enabled systems can be severe, potentially causing discrimination (J. A. Buolamwini, 2017; J. Buolamwini & Gebru, 2018; Dastin, 2022; Hoffmann & Söllner, 2014; Kayser-Bril, 2020; Olteanu et al., 2016; Ruiz, 2019) and even resulting in deaths (El Atillah, 2023; Kohli & Chadha, 2020; Pietsch, 2021).

Therefore, despite the advantages of AI, a phenomenon known as “algorithmic aversion persists”, causing consumers to hesitate in using AI even though algorithms frequently outperform human capabilities (Castelo et al., 2019; Dietvorst et al., 2015; Longoni et al., 2019;

Niszczota & Kaszás, 2020). Previous studies investigating AI acceptance have shown that consumers generally favor human labor over AI (Granulo et al., 2021) and harbor concerns regarding the ethical implications of AI (Ameen et al., 2021; Leung et al., 2018). Research by Fast & Horvitz (2017) shows that growing fear of AI is linked to concerns about losing control, job security, and the lack of moral judgment in the technology. Ensuring the trustworthiness of technology and its marketable products and services is therefore crucial, especially as AI becomes increasingly complex and inevitably ubiquitous in societies. (W. Li et al., 2023; H. Liu et al., 2023; Nickel et al., 2010; Petkovic, 2023). Due to its potential to revolutionize the way we work and live, this goes hand in hand with a focused effort required to ensure that the development of AI complies with ethical, legal, and social norms. (W. Li et al., 2023). Despite the substantial benefits AI offers, the potential societal risks underscore the importance of assessing its overall impact from a perspective rooted in public values (Medaglia, Gil-Garcia, & Pardo, 2021). As AI's development will continue regardless of individual approval or disapproval, the question is not whether, but how to shape this disruptive process (Brynjolfsson & McAfee, 2014).

However, how does one familiarize society with a technology that carries both significant risks and tremendous benefits? Within the scope of this research, this question will be addressed from a communication perspective: Media Richness Theory (MRT) will be employed to explore people's value systems by examining their perceptions of generative AI (GAI). By adopting this approach, the study aims to support AI systems designed with a strong emphasis on ethical considerations and effective communication of their trustworthiness. The goal is to identify best practices for these "positive" AIs to demonstrate their reliability, enabling users to differentiate between trustworthy and less reliable AI systems. By examining how these AIs convey their ethical commitments, the study seeks to enhance user confidence in AI technologies.

2.1.3 Technology Adoption and Marketing

As stated, the perception of end-users regarding AI systems is crucial for their successful implementation and acceptance (Vakkuri et al., 2020). An end-user is the person or group who uses a product, system, or service, particularly in IT. They interact with software, hardware, or

information systems to perform tasks or achieve goals, unlike developers or IT professionals who create or maintain these technologies. Understanding end-user needs is crucial because it ensures that information systems are designed to be relevant and useful for their intended job performance, ultimately leading to higher user satisfaction and system success (Mahmood et al., 2000). Understanding and positively influencing these perceptions can therefore significantly enhance the adoption of AI technologies (Ibáñez & Olmeda, 2022). An effective strategy to foster a healthy and sustainable development and adoption process of AI therefore involves strategic communication efforts that meet society's informational needs appropriately (Deloitte, 2022; The White House, 2023; van Wynsberghe, 2021). This approach aims to achieve that AI development is aligned with societal values, enhancing public understanding and trust.

Effective communication helps bridge the gap between AI advancements and public perception, promoting transparency and addressing ethical concerns (van Wynsberghe, 2021). Global advisory institutions emphasize that effective marketing strategies from AI development companies, governments, and other public sector organizations can significantly help meet informational needs by educating the public - This can involve public awareness campaigns, educational content, and proactive engagement with the community (Abillama et al., 2021; Berglind et al., 2022; Riparbelli, 2024). For instance, the World Economic Forum highlights the importance of public awareness campaigns to build trust and credibility around AI technologies (Riparbelli, 2024). These campaigns can include tutorials and positive use cases to show how AI can be leveraged beneficially. Similarly, McKinsey notes that governments can play a crucial role in educating the public about AI by implementing strategies that demonstrate AI's potential benefits while addressing concerns about its use (Berglind et al., 2022). This approach can lead to better public understanding and acceptance of AI technologies. The Boston Consulting Group emphasizes the need for AI education to ensure that the public can make informed decisions about AI applications (Abillama et al., 2021). Scientific research underscores this approach, stating that companies and institutions can enhance customer understanding of AI by providing clearer explanations of AI functionality, thereby increasing trust levels (Bedué & Fritzsche, 2022). Effective communication and education strategies can therefore help demystify AI and illustrate its practical benefits in various sectors, thereby fostering a more informed and receptive public.

As understanding societal perceptions, including their fears, hopes, and values, forms the foundation of successful marketing strategies (Rodriguez-Rad & Ramos-Hidalgo, 2018), it is essential to access and address these needs before spreading information, especially in the context of technology (Gordon, 2018; National Science Foundation, 2024; Yingprayoon, 2023). The key question is: how much information do people need? Do they want to know the technological details of the development process, or is it sufficient to assure them that safety measures are in place? Providing too much information might be overwhelming, while too little could lead to suspicion and distrust, similar to the negative effects of greenwashing in environmental protection (Reinhardt, 2023; Wright, 2024). Finding the right balance is crucial to ensure trustworthiness without causing confusion or skepticism (Wacksman, 2021).

To gain those insights into customer perceptions and make informed marketing decisions, it is essential to draw upon the in-depth knowledge provided by academic discourse (Kraus et al., 2023; S. Kumar et al., 2021). But even though AI sparked a new research focus in psychology and management, there remains a significant gap in understanding which characteristics of AI lead to aversion or acceptance (Jussupow et al., 2020). There has been limited emphasis on the end-user perspective of GAI and AI in general in the existing body of research (Degas et al., 2022). Most empirical studies have primarily focused on autonomous, self-driving vehicles and addressing data management requirements (Ahmad et al., 2023). Areas such as ethics, trust, and explainability require additional research (Ahmad et al., 2023). Since developers are eager to demonstrate that their technology functions, a lot of published studies on AI therefore currently concentrate on the technology itself, leaving unclear the effects of other factors on technology acceptance (Sujan et al., 2020). As predicting the acceptance of GAI from the perspective of the end-user is a relatively new topic, the research so far emphasizes the importance of exploring end-user perceptions and the variables that influence adoption (Ismatullaev & Kim, 2024a). A systematic mapping study conducted by Ahmad et al. (2023), focusing on end-user needs in AI, concludes that there is an urgent need for an approach that fully integrates all human-centered aspects in the development of AI-based software. The study found that the current discussions often focus more on generating ideas than on offering practical insights and explicit implementation strategies. There is limited understanding of how well theory aligns with real life. Therefore, the following section will first discuss the findings of these theories, which will then be examined in the context of the study.

One can summarize that due to rapid technological advancements, significant knowledge gaps exist in the current state of AI research (Ofosu-Ampong, 2024). However, what is already known, and what existing knowledge can be built upon and expanded?

First of all it is well-understood that positive user perception is one of the primary drivers that leads to success of technology (Al-Khaldi & Olusegun Wallace, 1999; Ditsa & MacGregor, 1995; Gelderman, 1998; Lyytinen, 1988; Schiffman et al., 1992; Szajna & Scamell, 1993). Ditsa & MacGregor (1995) identified several critical components influencing end-users' perception of technology: The quality of information , the user interface features, the support provided by the company (staff, manuals, etc.), the involvement of the user in the planning, development and implementation and the user attitudes towards the technology.

Research into consumer behavior and societal perception of AI, particularly generative AI, reveals a complex interplay of trust, transparency, and communication (Hoff & Bashir, 2015; Ismatullaev & Kim, 2024b; Shneiderman, 2020; B. Zhang & Dafoe, 2019). These factors significantly influence how AI is perceived and accepted by the public (B. Zhang & Dafoe, 2019). Trust is central to the acceptance of AI technologies (Hoff & Bashir, 2015; Ismatullaev & Kim, 2024b). It is multi-faceted, involving trust in the functionality of AI systems, ethical considerations in AI algorithms, and governance mechanisms overseeing AI applications (David et al., 2024). Technology trust therefore extends to the entities responsible for developing and regulating AI (Siegrist & Cvetkovich, 2000). Accordingly, the discrepancy between the current governance landscape of AI, largely driven by corporate self-regulation, and public expectations creates a trust deficit (David et al., 2024).

Effective communication strategies are crucial for building trust. Transparency in AI operations, such as clear explanations of AI recommendations, enhances user understanding and acceptance (Sinha & Swearingen, 2002). Transparent communication demystifies AI processes and engages users in a way that builds credibility (Sherman & Cohen, 2002). Additionally, Sinha and Swearingen (2002) emphasize how clear communication can influence user trust in recommender systems, suggesting that providing detailed explanations and rationale for AI recommendations can enhance user understanding and acceptance of AI technologies. Addressing biases within AI systems and ensuring fairness are significant to gaining public trust as bias can lead to unfair outcomes, which erodes trust (Mehrabi et al., 2022). Research highlights the need for rigorous testing and validation to demonstrate AI's

reliability and effectiveness (Lanus et al., 2021). By employing testing techniques, developers can enhance the quality and robustness of AI systems, ultimately influencing society's perception of AI reliability (Lanus et al., 2021). Engaging with users and incorporating their feedback into AI design can align AI systems with societal values and expectations (C. Wang et al., 2021) . This user-centered approach fosters a sense of collaboration and trust (Sinha & Swearingen, 2002; C. Wang et al., 2021). Selbst et al. (2019) emphasize the significant importance of fairness and accountability in AI-driven sociotechnical systems. Ensuring transparency, fairness, and accountability in AI decision-making processes is essential for fostering trust among end-users and society at large.Moreover, David et al. (2024) reveal the significant role of trust in shaping public attitudes towards AI. In the context of AI, trust is not only placed in the technology itself but also in the entities responsible for its development and regulation, including trust in the functionality of AI systems, trust in the ethical considerations embedded in AI algorithms, and trust in the governance mechanisms overseeing AI applications (David et al., 2024).

The psychological factors influencing user perceptions of AI are also crucial. Research by Sherman and Cohen (2002) explores the concept of self-affirmation in accepting threatening information, highlighting the psychological factors that influence user perceptions of technology. By acknowledging and addressing user concerns through open communication channels, organizations can reduce defensive biases and foster a more positive perception of AI technologies (Sherman & Cohen, 2002). This approach not only builds trust but also promotes user engagement and collaboration with AI systems (Sherman & Cohen, 2002).

In conclusion, discussions in both academic and public spheres often focus on establishing trustworthy AI through principles such as effectiveness, fairness, transparency, robustness, privacy, security, and human-centric values (Mittelstadt, 2019; Toreini et al., 2020; Varshney, 2019). While these principles address the technical aspects of AI trustworthiness (Liao & Sundar, 2022), trust itself is fundamentally a human judgment characterized by the perception of dependability in vulnerable situations (Lee & See, 2004; Vereschak et al., 2021). (Liao & Sundar, 2022) highlight that the perception of AI technology can vary among individuals, pointing out the importance of effectively communicating AI’s trustworthiness to users and stakeholders. Research indicates that transparency and effective communication are crucial for building trust in AI (David et al., 2024; Sherman & Cohen, 2002; Sinha &

Swearingen, 2002). Addressing biases and ensuring fairness in AI systems further enhances trust (Lanus et al., 2021). Engaging the public and incorporating their feedback is also vital for AI acceptance (Sinha & Swearingen, 2002). Strong governance, grounded in ethical frameworks, can bridge the trust gap (David et al., 2024). Afroogh et al. (2024) highlight that building trust in AI involves understanding factors related to AI itself, human interactions, and the specific context of use. Key technical factors like transparency, explainability, and performance are crucial for enhancing the AI system's trustworthiness across various domains. However, to achieve user trust, these attributes must be perceived as trustworthy, which can be facilitated through proper documentation and communication. Research demonstrates the need for effective communication strategies, yet the precise approach to enhancing public trust through information dissemination is still a topic for further investigation (Afroogh et al., 2024).

The societal impact of AI and the importance of public trust in its development cannot be overstated therefore. Addressing uncertainty, promoting transparency, mitigating bias, and ensuring rigorous validation processes are essential for enhancing trust and acceptance in AI technologies (David et al., 2024; Lanus et al., 2021; Sherman & Cohen, 2002; Sinha & Swearingen, 2002). As research progresses, exploring the relationship between communication strategies and consumer trust in GAI is crucial for developing effective information dissemination strategies and fostering a more positive perception of AI.

2.2 Understanding End-Users' Informational Needs

2.2.1 Media Richness Theory

Defining Media Richness: Media Richness Theory (MRT) provides a framework for understanding the effectiveness of different communication media in conveying information. The theory posits that richer media are more effective for complex, ambiguous tasks, while leaner media are suitable for straightforward, routine tasks (Daft & Lengel, 1986). The key dimensions of Media Richness that determine the appropriateness and effectiveness of a communication medium include immediate feedback, multiple cues, language variety, and personal focus. Media Richness theory (MRT), introduced by Daft and Lengel (1986) and later expanded with Trevino (1987) emerged in the 1980s alongside the rise of electronic

communication. The theory posits that communication effectiveness depends on matching the richness of the medium to the equivocality of the task. Per Daft and Lengel (1986), uncertainty is the absence of information, which can be alleviated by increasing the amount of information provided. Equivocality, however, refers to ambiguity or misunderstanding, which is better addressed by the richness or quality of the information. Accordingly, media with varying levels of richness have various communication effects (Daft & Lengel, 1986; Trevino et al., 1987).

Daft and Lengel's definition of Media Richness (MR), formulated before the advent of the electronic age, described it therefore as the "ability of information to change understanding within a time frame" (1986). Daft and Lengel's (1986) research on Media Richness theory primarily examined the factors influencing media choice. Hollingshead and McGrath (1993) later introduced the task-media fit hypothesis, which suggests that the effectiveness of a media choice depends on how well it matches the decision-making tasks. In subsequent research, Media Richness has been therefore defined as a set of objective characteristics - such as feedback capability, communication ability, language variety, and personal focus as previously worked out by Daft and Lengel (1986) - that determine how well a medium can convey rich information (Hollingshead & McGrath, 1993). The following will provide a closer explanation of these characteristics:

Immediate feedback refers to the ability of a communication medium to facilitate real-time, two-way interaction. This dimension is critical as it allows communicators to clarify misunderstandings, ask questions, and provide instant responses, thereby enhancing the accuracy and efficiency of the communication process. Daft and Lengel (1986) defined face-to-face communication as the richest medium in terms of immediate feedback because it allows for synchronous exchanges and nonverbal signals that aid in understanding.

Multiple cues encompass the variety of signals a medium can convey, such as visual, auditory, and verbal cues. Richer media can deliver a combination of these cues, which helps in transmitting nuanced and detailed information. Trevino et al. (1987) stated that face-to-face interactions provide visual cues like body language and facial expressions, as well as auditory cues such as tone of voice, which together enhance the clarity and emotional tone of the message. *Language variety* refers to the extent to which a medium supports the use of a wide range of language forms, from formal and technical language to informal and colloquial expressions. Daft and Lengel (1986) explained that rich media enable the use of varied language

that can convey subtle meanings and complex ideas effectively. This is particularly important in scenarios that require the expression of intricate concepts or emotional nuances. *Personal focus* pertains to the degree to which a medium allows for personalized and emotionally engaging communication. Carlson & Zmud (1999) indicated that richer media facilitate a personal connection between communicators, which can strengthen relationships and enhance the effectiveness of the message. This dimension is vital in contexts where trust and personal rapport are important, such as in management or counseling. *Task fit* is a dimension that considers how well a medium aligns with the specific requirements of the task at hand. Hollingshead & McGrath (1993) stated that communication media vary along a continuum of potential richness, and their effectiveness depends on the nature of the task. Tasks with high ambiguity and complexity benefit from richer media, while simpler, well-defined tasks are better suited to leaner media.

Scholars have since applied Media Richness theory in the context of marketing to study how media choice affects consumer decision-making (e.g. Lim & Benbasat, 2000; Maity et al., 2018; Maity & Dass, 2014). According to Cho et al. (2009), the theory's core principle is that effective communication depends on aligning the medium with the level of ambiguity, which requires a deep understanding of consumer needs. The key insight from Trevino et al. (1987) is that effective managers must align the medium’s richness with the task's equivocality to communicate effectively. Therefore, the communication process should consider the types of media tools used to ensure effective communication, emphasizing the importance of matching the media to consumer’s needs (task-media fit). As marketers strive to create compelling advertising content that captures and engages consumers (Shareef et al., 2017), it is essential for them to tailor their messages to meet the specific needs of their target audience. A comprehensive understanding of individual and societal perceptions, including fears, hopes, and values, is essential for the formulation of effective marketing strategies (Rodriguez-Rad & Ramos-Hidalgo, 2018). Consequently, it is of interest to marketers to understand the interplay between Media Richness and customer perception (Palvia et al., 2011). This can encompasses various dimensions such as the utilization of multiple communication cues or the use of natural language (Trevino et al., 1987).

Guerreiro et al. (2015) for example state that “the abundance and complexity of stimuli during purchase decisions may influence consumers” emotional and cognitive states, which

may trigger avoidance or approach responses.' MRT suggests that richer media, capable of providing more comprehensive information, can reduce uncertainty and equivocality (Daft & Lengel, 1986). Accordingly, Media Richness has been found to significantly affect trust and customer loyalty (Mohamad, 2020). The literature widely supports that Media Richness enhances advertising effectiveness by delivering more information (Lim & Benbasat, 2000; Suh, 1999; Treviño et al., 2000). Increased Media Richness reduces the cost of information search and expands the range of options consumers consider during decision-making (Maity et al., 2018a). The volume and accuracy of information for example improve customers' perception of quality in e-commerce (Song et al., 2012). McGrath and Hollingshead's (1993) task-media fit hypothesis suggests that media vary along a continuum of "potential richness of information" (Suh, 1999), with extremes being unsuitable for communication tasks due to either overwhelming richness or insufficient information. Thus, media with moderate richness are better suited for tasks involving choice and negotiation. Maity & Dass (2014) further explore this, finding that consumers prefer media with medium to high richness for complex decision-making, while simpler tasks are typically handled with lower richness media. Tseng & Wei (2020) found that Media Richness significantly affects the early stages of consumer behavior (attention, interest, and search) but has less impact on the later stages (action and sharing), suggesting that firms should use richer media to engage potential customers early in their buying journey. A study on AI-powered voice assistants highlights the importance of Media Richness in customer engagement across different cultures (P. H. Huang & Zhang, 2020). The study found that British users prefer low Media Richness for transactional purposes and high Media Richness for non-transactional purposes, while Taiwanese users show the opposite preferences. This suggests that AI voice assistant providers need to tailor Media Richness levels to user intentions and cultural contexts for effective engagement. The degree of Media Richness required is therefore context-dependent—greater Media Richness does not always result in a more positive marketing effect; in fact, it can sometimes lead to confusion or overwhelm.

Accordingly, the present study aims to evaluate the degree of Media Richness required in the context of GAI, a technology with such profound implications that it induces anxiety across individuals, society, and businesses on a global scale (Ferreira et al., 2022). Specifically, it will assess how different degrees of Media Richness in a text-based company statement document from a fictitious GAI company impact end-users' trust in both the technology and the

company. Company statements, also referred to as mission statements, are an effective approach to communicating a firm's core values to its stakeholders as Leuthesser & Kohli (2015) point out. Therefore, such statements are integral to the construction of corporate identity, which involves the articulation of an organization's underlying philosophy and strategic objectives to both internal and external audiences through communication, behavior, and symbolic representation. Mission statements are widely regarded as essential for defining a company's identity, purpose, and strategic direction, serving as a critical mechanism for transmitting the firm's core values to its stakeholders (Leuthesser & Kohli, 2015). Using a company statement to test trust levels within generative AI is more appropriate than using an advertisement within the scope of this study because company statements clearly and consistently reflect the organization's core values and long-term goals (Kaplan et al., 2010). They provide a more authentic, neutral perspective that appeals to a broad audience, making them a reliable measure of trust (Jo Hatch & Schultz, 2003). In contrast, advertisements are often persuasive and focused on short-term objectives, which may not accurately represent the company's foundational principles or foster trust in the same way (De Mooij, 2019). This approach aims to provide insights into the informational needs of end-users.

As a text-based company statement will be used to test trust levels, two dimensions emerge from the literature as relevant for this research. Building on Wang's (2022) research, which examined the crucial role of MR in user experience in a technological context, MR will be divided into two dimensions within the context of this study: Information content richness, which measures the extent to which the richness of a company statement document of a fictive GAI company's information content meets end-users' needs, and information quality richness, which refers to the quality and reliability of the information provided by the company to the end-users. Therefore, drawing on previous research, particularly the contributions of Trevino et al. (1987), Hollingshead & McGrath (1993), and Z. Wang (2022), this study defines Media Richness (MR) as *the extent and depth of information provided in text-based content relative to the effectiveness of a communication medium in meetingend end-users' informational needs.* This includes evaluating the richness of the information content to ensure it aligns with the audience's requirements and expectations, as well as assessing the quality and reliability of the information to ensure it is trustworthy and dependable for the intended recipients..

2.2.2 Interplay of Trust and Media Richness

Trust is a key factor in the relationship between Media Richness and communication effectiveness (Carlson & Zmud, 1999; Doney & Cannon, 1997; Fulk et al., 1987; Gefen et al., 2003; McKnight et al., 2002a; Trevino et al., 1987). At the same time, to successfully implement AI in business, industry, and society, a balanced system that relies on trust among users is essential (Bughin et al., 2018; Cheatham et al., 2019; Choung et al., 2023b; Floridi, 2007; Jacovi et al., 2021a; Shin, 2021; Taddeo & Floridi, 2018a). Trust is fundamental to how people see themselves, their surroundings and organize their lifes (Lewis & Weigert, 1985), and it can therefore act as a framework for understanding how users perceive GAI (Choung et al., 2023b; David et al., 2024). Trust is essential as a preliminary factor influencing individuals' decisions to engage in risk-taking behaviors or adopt new technologies (Chi et al., 2021; Jacovi et al., 2021b; Keding & Meissner, 2021). The possible harmful effects of using AI systems however have led to a lack of trust by users (Bach et al., 2024). End-users are more likely to embrace AI systems when they understand how these systems work, trust them to make reliable decisions, and recognize that trust is essential in determining how people interact with AI systems, driving acceptance and adoption (Alarcon et al., 2018). Trawnih et al. (2022) investigated how AI affects customer experiences in brand management, revealing that trust has the greatest impact on AI-powered interactions. An increasing number of researchers contend that building and preserving user trust is crucial for balancing the user-AI interaction (Jacovi et al., 2021a; Shin, 2021), creating reliable AI (Smith, 2019) and fully realizing AI's potential for society (Bughin et al., 2018; Cheatham et al., 2019; Floridi et al., 2018, Taddeo & Floridi, 2018a). Given that trust in AI is a key principle when addressing user needs, working with trust metrics is considered essential (Afroogh et al., 2024).

Similarly trust plays a crucial role within MR: The richness of the media through which brands communicate with consumers can influence perceptions of transparency, competence, and benevolence—key dimensions of trust (Afroogh et al., 2024; B. Li et al., 2023). Trevino et al. (1987) did not only demonstrate how different communication media can enhance communication effectiveness. They also highlighted that this enhancement is contingent upon the trust between communicating parties. In contexts where trust is lacking, such as facing as disruptive technology capabilities (Castelo et al., 2019; Dietvorst et al., 2015; Longoni et al.,

2019; Niszczota & Kaszás, 2020), the potential benefits of rich media may not be fully realized, suggesting that trust is integral to leveraging the full capabilities of Media Richness. Carlson & Zmud (1999) further explored the impact of Media Richness on decision-making processes, emphasizing the mediating role of trust. Their research indicated that trust in the communication medium significantly influences how effectively individuals use it. This finding suggests that Media Richness alone is insufficient for effective communication; trust must be established to fully capitalize on the benefits of richer media. This is particularly relevant for GAI, where user skepticism may impede the adoption and effective use of advanced communication technologies.

McKnight et al. (2002a) investigated trust in online environments, finding that it significantly affects how users perceive and utilize digital communication channels. Their study aligns with MRT principles, showing that trust enhances the effectiveness of richer media. In the context of GAI and the scope of this study, it is therefore of interest to evaluate the effectiveness of communication mediums and varying degrees of information provided with respect to trust.

Fulk et al. (1987) examined the interplay between communication technology and organizational effectiveness, highlighting the role of trust. They found that trust enhances the utility of richer media in organizational settings, facilitating more effective communication. This insight is critical for GAI, as organizations must build trust in AI-driven communication tools to ensure their successful implementation and utilization. Gefen et al. (2003) focused on trust in online transactions, finding that it impacts the perceived usefulness of communication media and users' willingness to engage with them. This reinforces the idea that trust is critical to effectively leveraging Media Richness. In the context of GAI, trust can enhance user engagement and satisfaction, leading to more positive outcomes in AI-driven communication. Doney & Cannon, 1997) explored trust in business relationships, finding that it facilitates more effective communication and helps overcome barriers associated with Media Richness. Their work underscores the importance of trust in enabling richer media to achieve their full potential, a concept that is highly relevant for the adoption of GAI technologies.

As the body of research shows, trust is crucial for effective communication, as well as for the successful integration of AI into today's and tomorrow's world. Consequently, this study will consider trust as a key determinant in accessing and addressing end-user needs in the context of this disruptive technology.

Understanding Trust: Trust can be described as an individual's readiness to rely on another party due to the qualities exhibited by that party (Rousseau et al., 1998). According to Mayer et al. (1995a), trust is defined as: 'The willingness of an individual to be vulnerable to the actions of another based on the expectations that the other party will perform a particular action that is important to the trustor, irrespective of the ability to monitor or control that other party'. In almost every kind of situation where there is uncertainty, interdependence or a possibility of unfavorable outcomes, trust is essential (Luhmann, 2018; Mayer et al., 1995b; Schneider & Fukuyama, 1996). Relationships, both personal and professional depend heavily on trust (Golembiewski & McConkie, 1975; Morgan & Hunt, 1994a). The development of mutuality and interdependence in human communication is inherently tied to the establishment of trust, thus traditionally associating the concept with interpersonal relationships (Choung et al., 2023b).

However, trust is a multidimensional construct that is connected to the traits, intentions, and actions of trustees (Lee & See, 2004; Schoorman et al., 2007). Numerous sources have noted that there are many different and unclear definitions of trust (e.g. Lewis & Weigert, 1985; Shapiro, 1987a). Psychology literature has long emphasized that trust is target-specific, meaning it involves one person trusting another with something valuable to them (LaFollette, 1996). Similarly, business researchers acknowledge the complexity of trust, arguing that the question "Do you trust them?" must be specified as "Trust them to do what?" (Mayer et al., 1995a). Building trustworthy relationships, therefore, also requires an understanding of how the individuals involved in the relationship themselves interpret the word (Fournier, 1998). Accordingly, the specific nature of trust implies that it should be analyzed from various perspectives (F. Li et al., 2008). Generally, studies on interpersonal and business trust view it as either a broad belief in the trustworthiness of another party (Moorman et al., 1992; Zucker, 1986) or as specific beliefs about the other party's integrity, benevolence, and competence (Larzelere & Huston, 1980; F. Li et al., 2008; Mayer et al., 1995b). An empirical study by Johnson-George and Swap (1982) found a general trust factor and two specific components—emotional trust and trust in a partner's reliability—highlighting that trust can exist at multiple levels.

According to David et al. (2024), there exists a correlation between trust in AI and the willingness to trust those involved in its development, with trust in AI systems and their

stakeholders being positively associated with the stakeholders' accountability in overseeing the technology. As this study focuses on understanding end-user perceptions of a technology product, it will examine trust levels through the lens of two sub-constructs: trust in the technology provider, referred to as *Brand Trust*, and trust in the technology itself, referred to as *Generative AI Trust*.

Defining Brand Trust: Brand Trust is a crucial element in marketing and consumer behavior, reflecting the extent to which consumers rely on a brand's promises and performance (Chaudhuri & Holbrook, 2001). This construct is multidimensional, encompassing both emotional and cognitive components that influence consumer perceptions, attitudes, and behaviors. Chaudhuri and Holbrook (2001) define Brand Trust as the degree to which consumers depend on a brand's ability to deliver on its promises, emphasizing the role of reliability in the brand-consumer relationship. (Munuera-Aleman et al., 2003) extend this definition by noting that Brand Trust involves consumer expectations of the brand's reliability and intentions, particularly in situations of risk or uncertainty. Trust, therefore, acts as a safeguard against perceived risks, boosting consumer confidence.

Li et al. (2008) identify two primary dimensions of Brand Trust: benevolence and competence. Benevolence relates to the brand's perceived goodwill and ethical behavior towards consumers, while competence refers to the brand's demonstrated ability to meet its promises consistently. For example, a brand engaged in CSR initiatives is perceived as more benevolent (Eisingerich & Bell, 2008), whereas a brand that consistently delivers high-quality products exemplifies competence (Morgan & Hunt, 1994).

Importantly, Brand Trust is not confined to commercial entities. It also applies to non-commercial institutions, including government agencies and non-profits. Trust in government institutions impacts public compliance with policies and civic engagement (Carter & Rogers, 2008), while non-profits rely on trust to achieve their social missions (Sargeant & Woodliffe, 2007). This broader view is especially relevant for emerging technologies like artificial intelligence (AI), where trust in both institutions and technologies is vital for public acceptance and adoption (David et al., 2024).

The influence of Brand Trust on consumer behavior is significant. It Brand Trustenhances customer loyalty, encourages repeat purchases, and fosters positive word-of-mouth (Delgado-Ballester & Alemán, 2005; Munuera-Aleman et al., 2003). Additionally, trust can mitigate the

effects of service failures, as consumers are more likely to forgive brands they trust (Reichheld & Schefter, 2000).

Theoretical frameworks such as Relationship Marketing Theory and Social Exchange Theory highlight the importance of trust in maintaining long-term customer relationships and facilitating mutually beneficial exchanges (Blau, 2017; Morgan & Hunt, 1994a). These frameworks underscore the multifaceted nature of Brand Trust and its role in positive consumer-brand interactions. (Gefen et al., 2003) explored the role of trust in online transactions, emphasizing that trust levels are a crucial determinant for effectively utilizing Media Richness. Their research demonstrated that higher Media Richness can enhance trust in the communication process, particularly regarding trust in brands and technology, suggesting a direct relationship between Media Richness and Brand Trust.

In this study, brand trust is thus conceptualized as a higher-level construct influenced by its various dimensions. Drawing on prior research by Li et al. brand trust is understood to arise when consumers place their confidence in a brand based on specific aspects such as performance competence and benevolent intentions. In this study, Brand Trust is thus conceptualized as a higher-level construct shaped by its various dimensions. Drawing on previous research by Li et al. (2008) and Jarvis et al. (2003), Brand Trust is posited to occur when consumers place confidence in specific aspects of a brand, such as its performance competence and benevolent intentions, with each dimension contributing to the overall perception of Brand Trust. Because this study aims to assess consumer perceptions, it is meaningful to examine Brand Trust through the lenses of competence and benevolence. As Li et al. (2008) argues, when consumers expect a brand to perform well, it can enhance their overall trust in the brand. However, trusting a brand's competence does not necessarily imply trust in its intentions or customer care. For instance, in a single purchase, customers may prioritize product performance over the seller's intentions. In contrast, long-term brand loyalty often involves trusting both a brand's competence and its benevolence towards customers. Understanding perceptions of competence and benevolence is therefore crucial for building sustainable relationships (Jarvis et al., 2003; F. Li et al., 2008).

Within the scope of this study, Brand Trust (BT) *is defined as consumers' " willingness to rely on the ability of the brand to perform its stated function,"* (Chaudhuri & Holbrook, 2001) *assessed through a two-dimensional framework. Here, competence is defined as the brand's*

ability to perform its stated function, while benevolence refers to the brand's reliability and intentions in situations involving risk to the consumer (F. Li et al., 2008).

It can be summarised that Media Richness refers to the ability of a communication medium to effectively convey information and meaning (Daft & Lengel, 1986). Studies indicate that higher Media Richness can enhance trust in the communication process, particularly in the context of Brand Trust (Gefen et al., 2003). Based on the media-task hypothesis (Hollingshead & McGrath, 1993), it is assumed that when users receive information about GAI through a rich media channels tailored to their informational needs, they are more likely to perceive the brand as reliable and trustworthy, thereby strengthening their overall trust in the brand. This assumption is supported by research indicating that richer media facilitate better communication and understanding, thereby enhancing trust (Carlson & Zmud, 1999). Thus, it is hypothesized that if the degree of Media Richness provided within a company statement by an institution working with GAI technologies meets the needs of its recipients, it can foster a deeper understanding and stronger emotional connection between the user and the GAI provider, leading to increased Brand Trust.

In this study, it is hypothesized that Media Richness is a predictor of Brand Trust.

H1: A direct relationship exists between Media Richness and Brand Trust

Defining GAI Trust: There are clear differences when it comes to trust in technology. The key distinctions between trust in technology and trust in humans lie in the emphasis on moral agency and competency for interpersonal trust, whereas trust in technology is rooted in functionality and reliability, underscoring the unique nature of trust in technology as opposed to trust in people (McKnight et al., 2011). Despite trust being extensively studied across various disciplines (McKnight & Chervany, 1996), there is still disagreement and insufficient research on the definition, importance, and measurement of user trust in AI-enabled systems (Bauer, 2019; Gulati et al., 2019; Sousa et al., 2019). Consequently, terms like trust, trustworthy, and untrustworthiness may be addressed without a clear focus and understanding of how these concepts influence interactions between users and AI systems, as mere principles cannot guarantee actual trustworthiness (Mittelstadt, 2019). Trust in technology is not easily quantifiable and cannot be simplified into a single construct (Choung et al., 2023). Instead, trust

in technology is a socio-technical concept that reflects an individual's willingness to be vulnerable to the actions of others, regardless of their ability to monitor or control these actions (Schoorman et al., 2007).

(Mayer et al., 1995) define trust in humans as encompassing beliefs in ability, benevolence, and integrity. Ability refers to competencies to successfully complete tasks, benevolence to positive intentions not solely self-serving, and integrity to consistent, predictable, and honest behavior. This three-dimensional view is crucial in human interactions, fostering reliability and integrity.

Research has extended this concept to human-technology relationships (Calhoun et al., 2019), particularly with technologies exhibiting human-like traits just as GAI is designed to do (Gillath et al., 2021). However, differences in user expectations and the non-applicability of interpersonal trust principles to human-machine trust have been noted (Madhavan & Wiegmann, 2007). Analyzing trust into its constitute elements allows for a deeper comprehension of technological acceptance (McKnight et al., 2002). Trusting is often based on beliefs and perceptions regarding the attributes and characteristics of the other party (X. Li et al., 2008). These beliefs and perceptions can be categorized into four dimensions: competence, benevolence, integrity (Callegaro et al., 2015; Iacobucci & Churchill, 2010; X. Li et al., 2008) and predictability (Iacobucci & Churchill, 2010; X. Li et al., 2008). McKnight et al. (2011) argue that trust in technology differs qualitatively from human trust, proposing dimensions of functionality, reliability, and helpfulness to replace ability, integrity, and benevolence, respectively.

AI, distinct from traditional technologies, operates autonomously and exhibits human-like characteristics, challenging existing trust definitions (Thiebes et al., 2021). For technologies with higher human resemblance, trust in human scales predicts outcomes better than trust in technology scales (Lankton et al., 2015), necessitating a conceptual framework that integrates technology and its human-like attributes. While a definitive definition of trust in AI is still evolving, two dimensions appear relevant: human-like trust (benevolence and integrity) and functionality trust (Mayer et al., 1995; Mcknight et al., 2011). Human-like trust concerns the alignment of AI algorithms with societal values and ethical design, while functionality trust pertains to reliability, competence, expertise, and robustness (Choung et al., 2023). Furthermore, AI-enabled systems are inherently complex, making it challenging for users to

immediately understand, accept, and justify their functions and design elements complex (Bathaee, 2018; Fainman, 2019; Mittelstadt, 2019). Even when users feel they have control over a complex system, they often misinterpret the causality of its elements (Dörner, 1980). These complexities clearly pose challenges in addressing user trust in AI systems and may lead to a trust gap between users and these systems (Ashoori & Weisz, 2019).

Trust in AI, as defined by Choung et al., involves having faith in three key aspects: benevolence, integrity, and competence (Choung et al., 2023). Benevolence and integrity pertain to human-like trust, emphasizing the kindness, sincerity, honesty, and consistency of AI technology. Competence, on the other hand, relates to functional trust, highlighting the effectiveness and reliability of AI systems. This definition will be adopted within the scope of this study, with *trust in Generative AI being defined as a multidimensional concept which includes human-like trust (benevolence and integrity), focusing on the social and cultural values embedded in GAI algorithms, and functional trust (competence), emphasizing the reliability and competency of GAI technology.*

Based on the media-task hypothesis (Hollingshead & McGrath, 1993), it is assumed that when users receive information about GAI through rich media channels tailored to their informational needs, it can foster a deeper understanding and a stronger emotional connection between the user and the AI technology. This, in turn, can lead to increased trust in generative AI. Research supports this notion, indicating that richer media enhance communication effectiveness and trust in technology (Carlson & Zmud, 1999). Therefore, it is hypothesized that if the degree of Media Richness in a Company Statement provided by an institution working with AI technologies meets the informational needs of its recipients, end-users are more likely to feel confident in the technology's benevolence, integrity, and competence, thereby increasing their trust in generative AI.

In this study, it is hypothesized that Media Richness is a predictor of trust in GAI.

H2: A direct relationship exists between Media Richness and Generative AI Trust

2.2.3 Value Systems

As previously noted, implementing GAI is not merely an organizational issue but also a moral one, prompting global society to consider: How do we want to shape the future? Moral

questions reflect deeply held values, which are crucial for understanding individuals' attitudes towards technology (Cruz-Cárdenas et al., 2019). Moreover, the potential societal risks of GAI underscore the need to assess its implementation process through the lens of public values (Medaglia et al., 2023). As (Choung ket al., 2023) further point out, trust in AI extends beyond its functionality to include the alignment of AI algorithms with societal values and ethical design. Understanding how individuals prioritize ethical values can shed light on their expectations in the AI domain (David et al., 2024). Addressing these value systems can influence trust in both the technology and its provider; when users perceive that technology aligns with their ethical and moral standards, their trust in both the brand and the technology increases (McKnight et al., 2002a). Embracing an AI approach based on human values benefits society and is essential for maintaining a positive brand image and ensuring long-term success in a dynamic market (Adomako & Nguyen, 2023; Hadj, 2020; Xie et al., 2024). Thus, examining value systems can help bridge the gap between how information is conveyed (Media Richness) and the resulting levels of trust. Consequently, understanding the interplay of value systems, Media Richness, and trust can provide valuable insights into individuals' informational needs and offer important clues for successful technology implementation. To achieve this, a deeper examination of the concept of "value" in the context of GAI is necessary.

Defining Values In an ethical or moral context, value refers to the principles or standards that individuals or societies use to judge what is right, important, or desirable. These values guide behavior, shape beliefs, and influence decisions, often reflecting deeper philosophical or cultural ideals (Rachels & Rachels, 2012). Values influence how individuals perceive and interpret information (Schwartz, 1992)., as well as their expectations and responses to communication (Schultz & Wehmeier, 2010). For instance, someone who highly values privacy might be likely to favor technology that offers strong data protection and transparent privacy policies. Companies, public organizations, and researchers have created numerous lists of AI ethics principles (Jobin et al., 2019). Given the absence of governmental regulations at the time of this research, these ethical guidelines currently can serve as guiding principles for organizations. They provide a framework for navigating the complexities of AI development and deployment, ensuring alignment with societal values and norms (Hagendorff, 2020).

This study aims to understand what customers anticipate from generative AI and how organizations offering these services can effectively support them to build trust in the

technology. Therefore, the criteria developed in ethical guidelines will serve as a foundation for categorizing various topics of potential customer concerns. The goal is to identify topics of importance for customers and assess whether different methods of explanation impact their trust.

AI Ethics Framework: AI ethics can be defined as a framework comprising values, principles, and methodologies that utilize universally recognized standards of morality to govern ethical behavior throughout the development and deployment of AI technologies (Leslie, 2019). (Attard-Frost et al., 2023) contextualize ethical guidelines within AI, offering recommendations for AI developers, users, policymakers, and other stakeholders to maximize benefits while minimizing associated risks. These guidelines are typically disseminated through research papers, reports, statements of principle, or similar documents, and are inherently prescriptive. They form an integral part of the broader AI ethics discourse, incorporating descriptive and critical perspectives from diverse sectors such as industry, academia, government, and civil society.

Ethical Guidelines and Customer Perceptions: Ethical guidelines can be used to assess customer perceptions because ethical considerations play a crucial role in shaping individuals' beliefs, attitudes, and behaviors towards various issues, including AI technologies (Mittelstadt, 2019). Customers tend to trust AI systems perceived as ethical and aligned with their values, highlighting the interconnected nature of ethics and trust in shaping customer perceptions (Allen & Wallach, 2012). By analyzing how individuals prioritize ethical values and principles within the AI context, research can glean insights into their perceptions of responsibility and trust towards various stakeholders involved in AI development and deployment (David et al., 2024). Ethical considerations offer a framework for evaluating the moral implications of technological advancements like AI, enabling individuals to assess potential risks and benefits (Mittelstadt, 2019). Understanding how individuals prioritize ethical values can shed light on their expectations in the AI domain (Choung et al., 2023; David et al., 2024).

Building Trust through Ethical AI Practices: Adhering to ethical standards signals companies' commitment to responsible AI practices, which can enhance customer trust and confidence in AI systems (Helberger et al., 2020). Research underscores the impact of ethical lapses in AI on customer trust, often resulting in erosion of trust following adverse events (Hoff & Bashir, 2015).Integrating ethical guidelines into AI development processes is essential for

mitigating risks and fostering sustainable trust with customers. Organizations that prioritize ethical considerations in their AI strategies are more likely to engender trust among customers, nurturing enduring relationships built on transparency and integrity (Arogyaswamy, 2020). Leveraging ethical guidelines to assess and uphold standards in AI technologies allows organizations to proactively address customer concerns and build a foundation of trust (Araujo et al., 2020). Ethical AI practices not only enhance organizational credibility but also contribute to establishing a trustworthy AI ecosystem that prioritizes customer and societal well-being.

Challenges and Criticisms of Ethical Guidelines: However, ethical guidelines should be reviewed critically. Stamboliev and Christiaens (2024) conducted a discourse analysis on ethics guidelines for trustworthy AI, noting that many proposals address ethical issues but are often too vague to be practically useful (Prem, 2023). Abstract principles require significant effort to be practically applied to AI systems, prompting calls to move from principles to practice (Jobin et al., 2019; Mökander et al., 2021). These frameworks lack specificity on application, leading to a range of approaches, methods, and tools being suggested to address AI ethics. As big tech's influence grows (Bartoletti, 2020), institutions striving for ethical standards often battle industry influence, leading to skepticism about the effectiveness of AI ethics (Munn, 2023). Critics warn that ethical principles in AI regulation often lead to non-binding statements, contributing to 'ethics washing' (Reinhardt, 2023; Wright, 2024). Despite over 80 AI ethics frameworks existing, conceptual vagueness and weak enforcement hinder ethics from countering industry interests (Mittelstadt, 2019). Against the backdrop of rapid technological advancements and fluctuating societal values, (Choung et al., 2023a) emphasize that these frameworks need to be adaptable and evolving. This is supported by (Afroogh et al., 2024; Zhou & Chen, 2023) who state that it's important to consider the rapid changes in AI development, which may sometimes clash with earlier principles or pre-set metrics. They suggest that AI should maintain flexibility within its intended framework, supported by data from questionnaires, interviews, and surveys, to keep up with public values amid the dynamic development process of AI (Afroogh et al., 2024).

The Scope of the Study: In line with the study's scope, since ethical guidelines are only used as a basis for capturing and categorizing potential customer perceptions, it is important to utilize a comprehensive category catalog. Ethical considerations surrounding AI encompass dimensions such as fairness, transparency, accountability, safety, and robustness (Araujo et al.,

2020). There is increasing agreement on the fundamentals of responsible AI development, although governance frameworks relating to AI systems and their widespread use are still developing (Choung et al., 2023a). Common ethical requirements for AI have been identified by Hagendorff (2020) through an analysis of ethics guidelines from organizations like Google, Microsoft, IBM, the OECD, IEEE, and governments of China, the United States, and the EU. These include human agency, privacy, transparency, explainability, safety, and cybersecurity. The European Commission's High-Level Expert Group on AI (AI HLEG) has provided a set of overarching principles that align with these concepts (David et al., 2024). Seven requirements emerged from their work: human agency, technical robustness, privacy and data governance, transparency, diversity and non-discrimination, social and environmental well-being, and accountability (see Figure 1). This framework is particularly relevant for this study as it connects the trustworthiness of AI with ethics and has been used to evaluate individuals' perceptions of trust regarding AI (Choung et al., 2023b; David et al., 2024).

***Figure 1**: Seven requirements for trustworthy AI proposed by European Commission*

Requirements	Description
1. Human agency and oversight	AI systems should allow people to make informed decisions. There should be a human oversight mechanism through a "human-in-the-loop" approach
2. Technical robustness and safety	AI I systems should be safe, reliable, and reproducible to minimize unintended harm
3. Privacy and data governance	Ensure privacy and data protection, which requires an adequate data governance framework
4.Transparency	AI systems and business models should be transparent, and the AI systems' decisions shouldbe explainable to the stakeholders. People need to be informed about the systems' capabilities and limitations
5. Diversity, non-discrimination, and fairness	AI systems should be accessible to all, and unfair biases should be avoided. Minimizing algorithmic bias is also important
6. Societal and environmental well-being	AI systems should benefit human beings and they should take into account the social impact and environmental consequences
7. Accountability	Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes

Notes. Source: David et al, 2024, p. 8

Based on the body of research presented, within the scope of this study, values are defined as *the ethical principles and standards that inform and guide the behavior, beliefs, and decision-making processes of individuals and organizations. These values reflect societal norms and cultural ideals, influencing the development, deployment, and acceptance of AI technologies. In the context of AI ethics, values encompass dimensions such as fairness, transparency, accountability, safety, and robustness, serving as a foundation for evaluating customer perceptions, fostering trust, and ensuring responsible AI practices.*

As shown, values significantly shape individuals' perceptions and attitudes towards technology. The interaction between Media Richness and values influences how information about technology is perceived and accepted (Suh, 1999). It is therefore hypothesized that a direct relationship exists between Media Richness and values. According to the task-media fit theory, it is consequently assumed that if the degree of Media Richness matches end-users'

informational needs, it can effectively convey ethical considerations and address users' value systems in the context of GAI.

In this study, it is hypothesized that Media Richness is a predictor of values.

H3: A direct relationship exists between the degree of Media Richness and values.

Additionally research shows that when users perceive that a technology aligns with their ethical and moral standards, their trust in both the brand and the technology itself increases (McKnight et al., 2002a). Values can therefore influence trust in both the provider and the technology, indicating that alignment with users' values fosters trust.
It is therefore hypothesized that there exists a direct relationship between values and Brand Trust, and values and Generative AI Trust, respectively.

In this study, it is hypothesized that values are a predictor of Brand Trust and GAI trust.

H4: A direct relationship exists between values and Brand Trust.

H5: A direct relationship exists between values and GAI trust.

Based on the research presented, it can be summarised and assumed that individuals' perceptions of the richness of a communication medium are shaped by their values. For instance, an individual who places a high value on transparency may find a detailed technological explanation of a system’s functionality both helpful and trust-enhancing, whereas another person might feel overwhelmed by the same level of detail. Therefore, understanding end-users' values can help institutions working with GAI tailor their communication to offer information that is both accessible and conducive to trust-building. The body of research indicates that understanding values can bridge the gap between how information is conveyed (Media Richness) and the resultant trust levels.

Accordingly, it is hypothesized that values mediate the relationship between Media Richness and Brand Trust, and Media Richness and Generative AI Trust, respectively.

H6: Values mediate the relationship between Media Richness and Brand Trust.

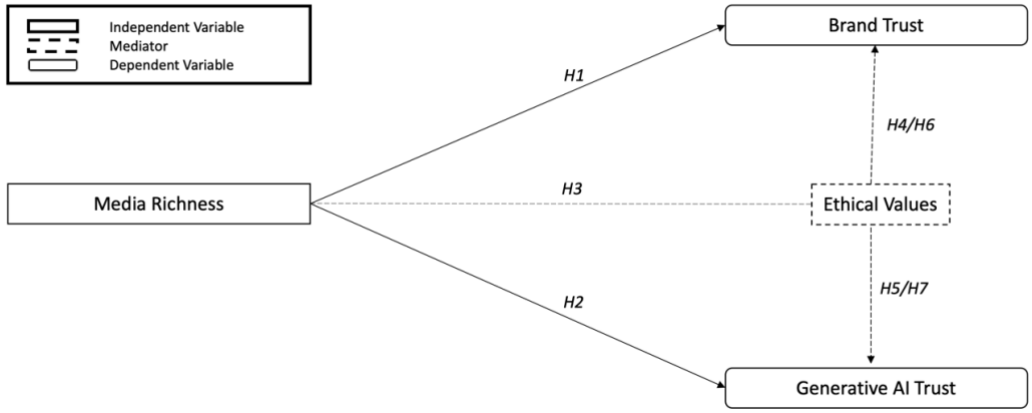
H7: Values mediate the relationship between Media Richness and Generative AI Trust.

This mediation effect suggests that Media Richness alone might be insufficient and reveals the extent to which the conveyed information must resonate with users' values to build trust effectively (Fulk et al., 1987).

2.2.4 Conceptual Framework

This study seeks to explore how the richness of Media Richness influences perceptions of trust within the scope of GAI. Figure 2 illustrates the conceptual framework outlining the research methodology. Media Richness is posited as a precursor to both Generative AI Trust (H1) and Brand Trust (H2). Additionally, it is hypothesized that Values moderate these relationships (H3, H4).

Figure 2 - *Conceptual Framework*



H1: A direct relationship exists between Media Richness and Brand Trust

H2: A direct relationship exists between Media Richness and Brand Trust

H3: A direct relationship exists between the degree of Media Richness and values.

H4: A direct relationship exists between values and Brand Trust.

H5: A direct relationship exists between values and GAI trust.

H6: Values mediate the relationship between Media Richness and Brand Trust.

H7: Values mediate the relationship between Media Richness and Generative AI Trust.

3 Methodology

3.1 Research Approach

This study seeks to identify patterns and draw broad conclusions about end-user informational needs around GAI by testing the proposed hypothesis grounded in the existing literature. Therefore, a quantitative research approach will be employed to facilitate quantitative predictions. Additionally, quantitative research permits the collection of data from a larger sample, measurement of data, generalization of findings, and identification of patterns (Malhotra, 2007). Given that the unit of analysis is consumer trust the questionnaire survey method has been selected to test the proposed hypotheses (Wood, 2024). To test the hypotheses, an experiment was designed involving one independent variableMedia Richness with three distinct levels. The manipulated independent variable was media richness, categorized as low, moderate, and highMedia Richness. As a result, three experimental groups were surveyed online. Each participant was randomly assigned to one test group. Each group answered an identical set of questions to measure Brand Trust, GAI Trust and Values, differing only in their prior exposure to different levels of Media Richness.

3.2 Stimuli

This study aims to investigate end-users' informational needs regarding generative AI and how organizations can cultivate trust in this technology. To operationalise MR, a company statement from a fictional generative AI service provider with varying levels of MR is designed (see Annex H). In developing the stimuli for this study, not all dimensions of Media Richness could be effectively represented across different communication mediums. Given that company statements are a widely recognized and practical form of communication (Leuthesser & Kohli, 2015), they were selected as the focus for this research. Company statements are easily implemented in practice, making them a suitable choice for the study. However, it is important to note that certain MR dimensions, such as immediate feedback, cannot be fully captured within the scope of company statements. Therefore, the study focused on adapting only the applicable dimensions of MR that could be operationalized in this medium—specifically, the

extent and depth of text-based information, as identified by (Z. Wang, 2022). This approach is consistent with previous research where studies (e.g., Carlson & Zmud, 1999; Sen et al., 2009; Suh, 1999) similarly adapted MR dimensions to fit the constraints of the medium being studied, based upon the task-media-fit hypothesis (Hollingshead & McGrath, 1993). Building on the research by (David et al., 2024), which evaluated end-user perceptions within the ethical guidelines set by the European Commission, this study employs the same framework to categorize potential customer Values, thereby informing the design of study stimuli. Additionally, this framework will examine whether participants' Values mediate their levels of trust in generative AI. Specifically, the study will present three company statements reflecting low, moderate, and high degrees of Media Richness. Within this context, Media Richness refers to the extent and depth of information provided in text-based content:

(1) *Low Media Richness:* This document presents the company's measures for trustworthy GAI implementation, solely communicating *what* the company is doing for trustworthy generative AI. The document is two pages long, consisting of a cover letter and a second page with an enumeration of the seven values of the company.
(2) *Moderate Media Richness:* This document adds the *why* to the *what*, educating end-users about what these measures mean in the context of generative AI and why they are important. The document is six pages long, including the same cover letter and enumeration of the seven values (*the what*), followed by short paragraphs explaining why these measures are important (*the why*).
(3) *High Media Richness:* This document includes the *how* in addition to the *what* and *why*, providing detailed technical explanations. The document is nine pages long, consisting of the same cover letter and enumeration of the seven values (*the what*), followed by a single page for each value with explanations of why these measures are important (*the why*) and detailed technical implementation descriptions (*the how*).

By varying the level of Media Richness, this study seeks to determine the optimal amount of information needed to foster trust in generative AI among end-users. The technical information within the document was developed in direct collaboration with the generative AI company Ella Tec GmbH, and the document's design was aligned with the established design standards

of leading AI firms (e.g., Adobe, 2024; Meta, 2021; Microsoft Corporation, 2022). To ensure that the stimuli accurately reflected the intended MR dimensions, a pre-study was conducted, followed by a manipulation check in the main study. These steps were crucial in verifying that the stimuli effectively represented the targeted dimensions of MR and achieved the study's objectives.

3.3 Data Measurement and Scales

The questions in the questionnaire were designed using scales from the literature to measure each variable in the model. These items were adapted from prior research, utilizing scales that have been extensively tested and validated, thereby establishing a robust measurement foundation. Figure 4 below displays the number of items for each scale and links each variable to its respective scale's author.

Figure 3 *- Scales authors and number of items*

Variable	Scale's Author	Na of items
Media Richness	Z. Wang (2022)	6
Generative AI Trust	David et al. (2024)	11
Brand Trust	F. Li et al., 2008	9
Values	Choung et al., 2023a	7

All items in this study were rated on a 5-point Likert scale, with options ranging from '1 = Strongly Disagree' to '5 = Strongly Agree' for MR, GT, and BT, and '1 = Not at all important' to '5 = Very Important' for Values (see Annex B). MR was evaluated using 6 items, GT with 11 items, BT with 9 items. In the final segment of the survey, participants were asked about their values regarding GAI, using 7 items.

Media Richness: To measure Media Richness, the study uses the scales of Information Content Richness and Information Quality Richness from (Z. Wang, 2022) because they are specifically relevant for evaluating text-based content. *Information Content Richness* ensures the text is comprehensive and meets user needs, while *Information Quality Richness* evaluates

the accuracy and reliability of the information. These dimensions are essential for assessing the effectiveness of a Company Statement focused solely on text, making them more appropriate than scales designed for multi-media contexts.

Brand Trust: (F. Li et al., 2008) developed scales to measure Brand Trust, providing a robust framework for evaluating trust in technology companies. Based on established trust theories (Mayer et al., 1995c; Singh & Sirdeshmukh, 2000), these scales focus on competence and benevolence. Competence assesses a company's ability to deliver effective solutions, while benevolence reflects its commitment to customer interests and ethical practices (Y. D. Wang & Emurian, 2005). These dimensions are crucial for evaluating trust in technology firms, as they ensure reliability and a user-centric approach to tech products and services (McKnight et al., 2002a).

Generative AI Trust: To measure GAI Trust, the study uses measures based on the work of Choung et al. (2023a), who employ a format adapted from (Mcknight et al., 2011). AI Trust assessed using a structured questionnaire, where participants express their level of agreement or disagreement with various statements regarding AI. This approach allows for the assessment of two distinct dimensions of trust: trust in human-like attributes of AI (HTAI) and trust in the functionality of AI (FTAI). The scales are effective for assessing trust in generative AI because they capture both human-like interactions and specific functionalities of generative AI, are validated through confirmatory factor analysis (CFA), address contemporary concerns of evolving AI technologies, and incorporate both cognitive and emotional aspects of trust. These factors make the scales valuable for evaluating trust in generative AI, supporting the development of effective use policies and governance.

Values: The European Commission's AI HLEG principles form a robust framework linking AI trustworthiness with ethical standards, influencing customer perceptions of AI trust (Choung et al., 2023a; David et al., 2024). Building on this foundation, Choung et al., 2023a developed scales to measure people's values in AI contexts, inspired by the European Commission's initiatives. These scales offer a structured approach tailored to AI values, enabling a thorough assessment of customer perspectives.

3.4 Data Collection and Sample

3.4.1 Questionnaire Development

For the purpose of collecting data and implementing the study, an online questionnaire was developed in English (see Annex A). Since this research focuses on consumer perceptions, a questionnaire was selected based on previous studies that assessed perceptions related to AI (Choung et al., 2023a; David et al., 2024). The online survey tool Tivian Unipark (EFS survey by Tivian, Tivian XI GmbH, Cologne, Germany) was used to create the questionnaire as well as for data collection and monitoring of the study. The survey comprised four consecutive parts: presentation of stimuli (1), measurement of MR, GT, and BT (2), assessment of Values (3), and the collection of socio-demographic characteristics (4).

At the beginning of the questionnaire, participants were informed that their participation was voluntary and could be terminated at any time. They were assured that their data would be treated anonymously and confidentially, and they had to give consent for the processing of personal data. Next, they were informed about the aim of the study. Participants were informed and encouraged to provide thoughtful responses to the survey, as commitment positively impacts data quality, including response accuracy (Hibben et al., 2022). The attention check scales were included based on a study by (Kung et al., 2018).

The study proceeded as follows: (1) After receiving a definition of generative AI, to prevent difficulties individuals face in understanding the technology and to be able to compare the results with existing literature (Kelly et al., 2023; Liang & Lee, 2017), participants were automatically and randomly assigned to one of three experimental groups: low, medium, or high Media Richness, according to the research design. (2) Participants were then asked in detail about their perception of MR, GT, and BT to evaluate the Company Statement. (3) Next, participants were presented with a list of values concerning the development and design of GAI technologies and were asked to rate the importance of each. (4) The final part of the questionnaire consisted of questions about the basic socio-demographic profile of the participants to characterize the sample.

3.4.2 Sample

The questionnaire was forwarded to friends and family and acquaintances randomly and with no specific requirements, meaning it was a convenience sample. The participants were recruited over a period of 60 days via various online social networks. Participation was accessible via desktop as well as via mobile devices. In total, completion took 8 minutes on average.

All data collected from the questionnaire across the three experimental groups were uploaded directly into IBM SPSS Statistics 29 and cleaned. This process involved removing all persons who have refused consent ($N = 29$), did not pass the three attention checks (N = 37), were under 18 years old ($N = 0$), or did not complete the whole survey ($N = 111$), were excluded. This resulted in a sample of 300 valid answers out of 477 questionnaires received, which were included in the analysis. Thus, the response rate was 62.89 %. The cleaned dataset was used to run descriptive Statistics analysis in SPSS and imported into SmartPLS 4 for inferential analysis using partial least squares structural equation modeling (PLS-SEM) to test the proposed model.

3.4.3 Pre-test of Constructs

To ensure the reliability and validity of the constructs prior to the large-scale sample collection, a preliminary study was conducted (see Annex C). This step aimed to mitigate the risk of unreliable results in the main study. The pilot test assessed the questionnaire's readiness for full implementation by identifying unclear concepts, resolving doubts about specific questions or topics, evaluating comprehension of the Company Statement, and eliminating redundant questions. An online questionnaire was administered with a sample size of $N = 21$, which was independent of the main study. No suggestions, doubts, or criticisms were raised by the respondents.

In the preliminary study, participants were shown only one version of the company statement, specifically the version with the highest degree of Media Richness. Following the presentation of the document, participants were asked to evaluate the constructs, mirroring the procedure designed in the main study (see Data Collection and Sample). The questionnaire also

included sociodemographic characteristics in the second section. The procedure was consistent for all participants (see Annex A).

To evaluate the internal consistency of the MR, GT, and BT constructs, Cronbach's alpha coefficients were calculated, revealing satisfactory levels of internal consistency, with all values exceeding the widely accepted threshold of .70 (J. F. Hair et al., 2010). Specifically, Cronbach's alpha values were 0.876 for MR, 0.936 for GT, and 0.843 for BT. To further assess the reliability of these constructs, composite reliability (ρc) was calculated, which also indicated high reliability for all constructs: 0.908 for MR, 0.944 for GT, and 0.877 for BT. These values demonstrate that the constructs exhibit strong internal consistency and are reliable measures.In addition, the average variance extracted (AVE) was computed to evaluate the convergent validity of the constructs. An AVE of 0.50 or higher is typically considered acceptable, as it indicates that the construct explains at least 50% of the variance in its indicators. The AVE values for MR (0.628) and GT (0.629) exceeded this threshold, suggesting satisfactory convergent validity. However, the AVE for BT was slightly below the recommended level, at 0.461. While this value suggests that less than 50% of the variance in the BT indicators is explained by the construct, it may be considered acceptable in the context of this study, given the relatively small sample size (N = 21).

4 Results and discussion

4.1 Overview

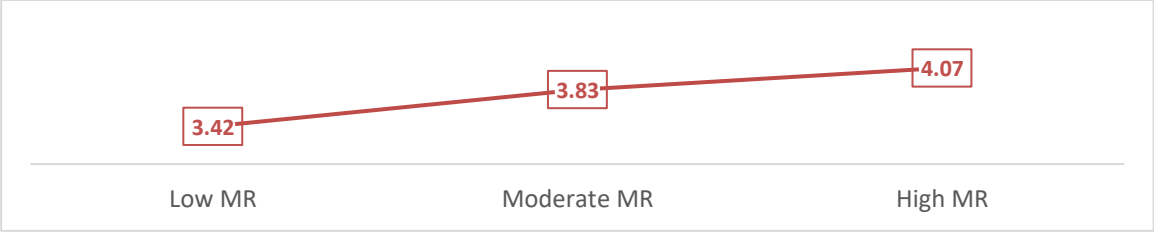
The study first uses IBM SPSS Statistics Version 29 for a descriptive overview of the data. In the second step, it employs partial least squares structural equation modeling (PLS-SEM) using SmartPLS 4 for inferential statistical analysis. This research evaluates the research model through two stages: the outer model (measurement model) and the inner model (structural model) (Henseler et al., 2016). Hypotheses were tested using bootstrapping resampling with 5,000 samples.

4.2 Preliminary control checks

4.2.1 Stimuli Perception

Since the stimuli were self-generated, a manipulation check was necessary to confirm that the manipulation had the intended effect within the experiment's parameters. Specifically, the manipulation check assessed whether the stimuli were perceived as intended, ensuring that the variable was measuring what it was supposed to measure. To achieve this, a preliminary assessment was conducted with the three test groups (N = 300) to establish a baseline for evaluating how different levels of Media Richness (low, moderate, high) influenced the dependent variables (Values, BT, and GT). This initial check helps assess the content validity of the stimuli and the effectiveness of the usage policy document in conveying the intended message. Subsequently, an Analysis of Variance (ANOVA) was performed to determine if there are statistically significant differences in perceived Media Richness among the test groups (see Annex E).

Figure 4: Perceived Media Richness Mean scores across groups



Note: Scale ranges from (1) to (5).

Descriptive statistics revealed that the mean perceived Media Richness scores were as follows: Low Media Richness (M = 3.418, SD = 0.813), moderate Media Richness (M = 3.825, SD = .678), and high Media Richness (M = 4.068, SD = .555). These results indicate the intended trend of increasing perceived Media Richness from the low to high groupsMedia Richness. The results of ANOVA showed a significant effect of Media Richness on Media Richnessperceived effectiveness, with an F-value of 22.322 and a p-value < 0.001, indicating that differences exist among the groups. Post hoc analysis using Tukey's Honestly Significant Difference (HSD) test revealed that the low Media Richness group (Mean = 3.418) was significantly different from both moderate Media Richness (Mean = 3.825) and high Media Richness (Mean = 4.068).

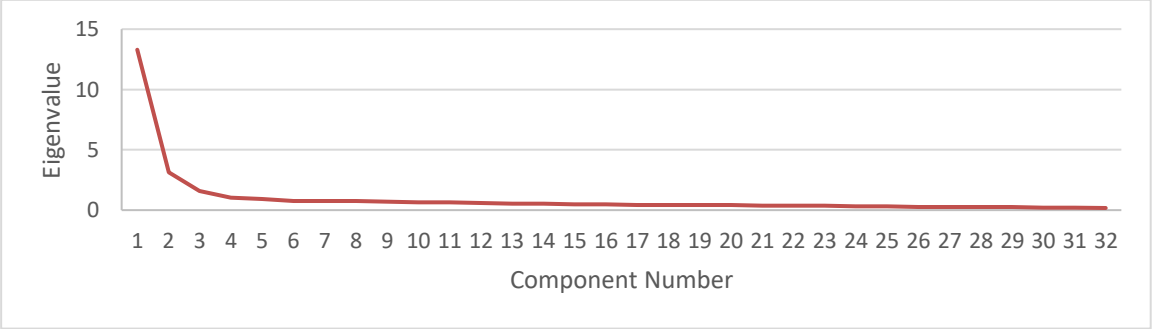
Thus, The intended effect was successfully communicated to the participants, with significant results demonstrating that the levels of Media Richness were perceived as progressively increasing between the groups. This confirms the content validity of the stimuli, allowing for further analysis.

4.2.2 Common method bias

Common method bias can be a concern (Podsakoff et al., 2003) when gathering behavioral and attitudinal data using self-report questionnaires at a single time point (Chang, Van Witteloostuijn, & Eden, 2010). Common method bias refers to the distortion in research results that occurs when the measurement method, rather than the constructs being measured, accounts for a significant amount of the variance observed, often due to collecting data from the same

source or using the same measurement technique. To mitigate this bias, the study ensures that all participants are aware of the survey's confidentiality, thereby reducing social desirability bias (Ganster et al., 1983; Podsakoff et al., 2003). Additionally, participants had to agree to provide thoughtful answers, as this transparency enhances data quality and response accuracy (Gummer et al., 2021; Hibben et al., 2022). To assess the potential for common method bias statistically, Harman's Single-Factor Test was conducted using exploratory factor analysis (EFA). The results indicated that the first factor accounted for 41.548% of the total variance (see Annex E). Since this is below the 50% threshold, common method bias is unlikely to be a significant concern in this study. Additionally, the scree plot (see Figure 5) supports this conclusion. The plot shows a clear "elbow" after the first factor, with the eigenvalues leveling off, indicating that no single factor dominates the variance.

Figure 5 - *Harman's Single-Factor Test – Component Analysis Scree Plot*



The scree plot further confirms that the first factor does not account for the majority of the variance. The "elbow" point at the second factor suggests that additional factors contribute to the variance, which mitigates the risk of common method bias. According to Podsakoff et al. (2003), if the first factor explains less than 50% of the variance, common method bias is considered to be less problematic. Therefore, the findings from both the total variance explained table and the scree plot indicate that common method bias is not a significant concern in this study.

4.3 Descriptive Statistics

4.3.1 Socio-demographic characteristics

Table 1 provides an overview of the socio-demographic characteristics of the sample. Detailed illustrations of each classification question can be found in Appendix X. Among the valid questionnaires, 51.3 % (N = 154) of the respondents were men, 37.0 % (N = 141) were women, 1.3 % (N = 4) identified as diverse, and 0.3 % (N = 1) did not specify their gender. The respondents were on average 34,54 years old (*SD* = 10.42). The sample showed a tendency towards a higher level of education, with 79,9 % of the participants (*N* = 240) having an attained a Bachelor's degree or higher qualification. 38.6 % (*N* = 116) had an average net income of less than 500 € per month, 27.0 % (*N* = 82) had an income of 500 – 1499 €, while 34 % (*N* = 102) had 1500 € or more at their disposal. The sample is not heterogeneous with regard to the current situation; the majority of participants, comprising 61.3% (*N* = 184), were workers.

Table 1 *- Socio-demographic characteristics of the sample*

		n	%	M	SD
Age				34.55	10.42
Gender	Male	154	51.3		
	Female	141	47.0		
	Diverse	4	1.3		
	Do not specify	1	0.3		
Education	No schooling completed	0	0		
	Basic education. (9th grade)	4	1.3		
	High school (12th grade)	45	15.0		
	Bachelor's degree	154	51.3		
	Post Graduation	25	8.3		
	Master's Degree	55	18.3		
	PhD	6	2		
	Other situation	24	8		
	Do not specify	2	0.7		
Income	No income	24	8.0		
	0 – 249 €	61	30.3		
	250 – 499 €	31	10.3		
	500 - 999 €	43	14.3		
	1000 – 1499 €	39	13.0		
	1500 – 1999 €	23	7.7		
	2000 – 2499 €	34	11.3		
	2500 – 2999 €	22	7.3		
	3000 € or more	23	7.7		
Current Situation	Student	21	7.0		
	Working Student	28	9.3		
	Worker	184	61.3		
	Unemployed	38	12.7		
	Other situation	24	8.0		
	Do not specify	2	0.7		

Notes. N = 300

4.3.2 Variable Metrics

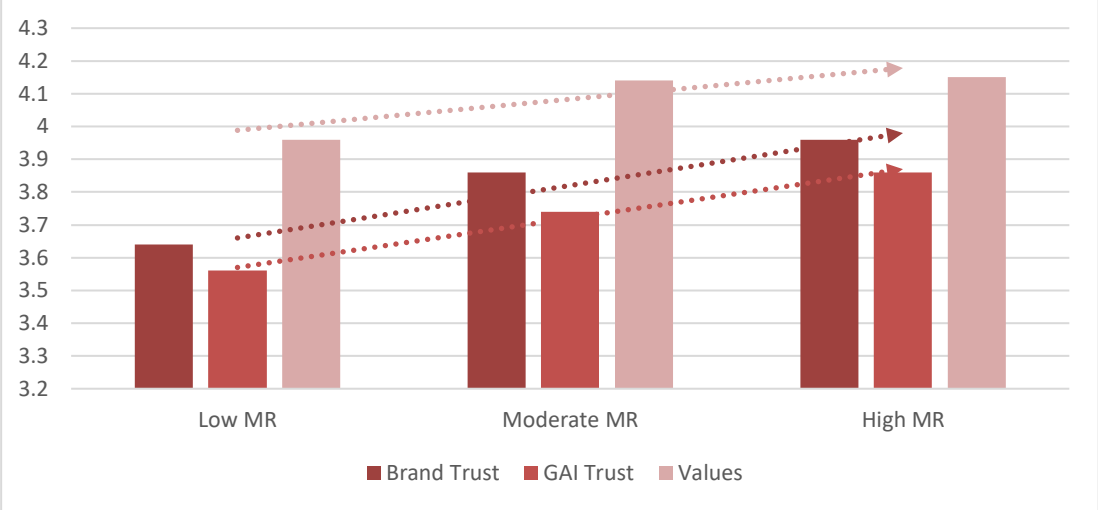
To analyze the descriptive statistics for the key variables under study, the descriptive statistics and frequencies functions of SPSS were utilized. This procedure enabled the calculation of essential measures of central tendency, variability, and frequency distributions, thereby offering a comprehensive overview of the data. Table 2 presents a detailed overview over the mean scores of the dependent variables, according to the exposition of Media Richness. Figure 7 illustrates the pattern of increasing mean scores for Brand Trust, GAI Trust, and Values with higher levels of Media Richness.

Table 2 *- Mean scores of dependent variables by Media Richness Levels*

Variable	MR Levels	*M*	*SD*	*n*
BT - Brand Trust	(1) Low	3.64	.71	102
	(2) Moderate	3.86	.60	102
	(3) High	3.96	.61	96
	Total	3.82	.66	300
GT – Generative AI Trust	(1) Low	3.56	.67	102
	(2) Moderate	3.74	.63	102
	(3) High	3.86	.57	96
	Total	3.71	.63	300
Values	(1) Low	3.96	.69	102
	(2) Moderate	4.14	.64	102
	(3) High	4.15	.63	96
	Total	4.08	.66	300

Notes: Scale ranges from (1) to (5). n = Valid Responses

Figure 6 - *Mean scores of dependent variables by Media Richness Levels*



Note: Scale ranges from (1) to (5). Lines present Linear Trendline.

Brand Trust: The mean scores for Brand Trust increase with higher levels of Media Richness. Participants exposed to high Media Richness (*M* = 3.96) reported the highest Brand Trust, followed by those exposed to moderate (*M* = 3.86) and low Media Richness (*M* = 3.64). The overall mean for all groups combined is 3.82, with a standard deviation of 0.66, indicating moderate variability in Brand Trust scores.

GAI Trust: The mean scores for GAI Trust also show an increasing trend with higher Media Richness. Participants exposed to high Media Richness (*M* = 3.86) reported the highest GAI Trust, followed by those exposed to moderate (*M* = 3.74) and low Media Richness (*M* = 3.56). The overall mean for GAI Trust is 3.71, with a standard deviation of 0.63, suggesting slightly lower variability compared to Brand Trust scores.

Values: The Values scores are generally higher than both Brand Trust and GAI Trust scores across all levels of Media Richness. The mean scores increase slightly from low Media Richness (*M* = 3.96) to high Media Richness (*M* = 4.15), with an overall mean of 4.08 and a standard deviation of 0.66, indicating similar variability to Brand Trust scores.

There is a consistent pattern of increasing mean scores for Brand Trust, GAI Trust, and Values with higher levels of Media Richness. The relatively low standard deviations indicate that the scores are closely clustered around the means, suggesting consistent responses within

each media richness group. The overall higher mean scores for Values suggest that participants generally rated the Values construct more positively compared to BT and GT. Given that the Values construct, acting as the mediator, exhibits relatively high mean scores on a Likert scale with a maximum value of 5, and that these mean scores show only a minor increase from low to high Media Richness (MR), it is valuable to conduct a more detailed analysis of this construct within the scope of the research. Since personal value systems are deeply intertwined with individuals' overall self-perception and their view of the external world, it is crucial to identify which specific values are most significant to participants. Understanding these values can offer valuable insights into how they influence perceptions and interactions with GAI. Table 3 provides insight into participants' value system concerning generative AI, based on the mean scores and standard deviations of the individual items within the Values construct.

Table 3 *- Mean scores of Value Items*

Item	M	SD	n
(1) Accountability: AI at any step is accountable for considering the system^s impact in the world.	3.96	.94	300
(2) Transparency: Transparency requirements that reduce the opacity of systems.	4.11	.92	300
(3) Privacy and data governance: Competent authorities who implement legal frameworks and guidelines for testing and certification of AI-enabled products and services.	4.22	.94	300
(4) Diversity, non-discrimination and fairness: The application of rules designed to protect fundamental human rights, such as equality.	4.07	1.03	300
(5) Technical robustness and safety: Systems are developed in a responsible manner with proper consideration of risks.	4.21	.83	300
(6) Human agency and oversight: Human oversight and control throughout the lifecycle of AI products.	4.07	.93	300
(7) Societal and environmental well-being: AI systems that conform to the best standards of sustainability and address like issues climate change and environmental justice.	3.92	.98	300

Notes: Scale ranges from (1) to (5). n = Valid Responses.

Overall Value System: The mean scores for all items are relatively high, ranging from 3.92 to 4.22, indicating that participants generally place a strong emphasis on these values when it comes to generative AI. The values related to Privacy and Data Governance, Technical Robustness and Safety, and Transparency receive the highest means, suggesting that participants prioritize these aspects most when evaluating AI systems.

Areas of Greater Emphasis: Privacy and Data Governance (M = 4.22) and Technical Robustness and Safety (M = 4.21) have the highest mean scores. This suggests that participants place significant importance on ensuring AI systems are safe, well-regulated, and adhere to

privacy standards. Transparency (*M* = 4.11) also ranks high, indicating a strong preference for reducing opacity and enhancing clarity in AI systems.

Values with Slightly Lower Scores: Accountability (*M* = 3.96) and Societal and Environmental Well-being (*M* = 3.92) have the lowest mean scores among the values but are still relatively high. This suggests that while these aspects are valued, they may be considered slightly less critical than the other factors. The slightly lower emphasis on Societal and Environmental Well-being compared to other values could indicate that while participants acknowledge the importance of sustainability and societal impact, it may not be their top priority in the context of AI systems.

Variability: The standard deviations are relatively close across the items, ranging from 0.83 to 1.03, indicating a moderate level of agreement among participants about the importance of these values. The smallest standard deviation is for Technical Robustness and Safety (*SD* = 0.83), suggesting a higher level of consensus about its importance.

Participants generally view all the values related to generative AI as important. Although all values are rated highly, there is a slightly lower emphasis on Accountability and Societal and Environmental Well-being. The relatively consistent standard deviations suggest a generally shared perspective among participants, with some variation in the degree of emphasis placed on different values.

4.4 Inference Statistics

4.4.1 Outer Model Results

Reflective measurement model: The research examines three facets to assess the reflective measurement model: convergent validity, internal consistency reliability, and discriminant validity. The detailed findings are shown in Table 4. The study removes two indicators (GT10, GT11) for the GAI Trust construct from the original model due to their low outer loadings. All the outer loadings in the seven reflective measurement models are at least 0.7 (J. F. Hair et al., 2010), varying from 0.7 to 0.89, and are statistically significant ($p < 0.001$).

For internal consistency reliability, Cronbach's alphas and composite reliabilities for all constructs are above the required 0.70 (Nunnally, 1978) and 0.60 (Bagozzi & Yi, 1988)

respectively, ranging from 0.82 to 0.92. The average variances extracted (AVEs) for all constructs range from 0.51 to 0.62, all exceeding the threshold of 0.5 (J. F. Hair et al., 2010), indicating high reliability of the indicators. These results demonstrate that the models are internally reliable.

Table 4 *- Reliability and validity test for the complete data*

Constructs	Indicators	Outer loadings	α	CR	AVE
Media Richness	MR1	0.73	0.88	0.91	0.62
	MR2	0.75			
	MR3	0.79			
	MR4	0.89			
	MR5	0.8			
	MR6	0.75			
GAI Trust	GT1	0.81	0.91	0.93	0.59
	GT2	0.73			
	GT3	0.76			
	GT4	0.77			
	GT5	0.74			
	GT6	0.8			
	GT7	0.7			
	GT8	0.74			
	GT9	0.8			
Brand Trust	BT1	0.8	0.92	0.94	0.62
	BT2	0.71			
	BT3	0.8			
	BT4	0.8			
	BT5	0.8			
	BT6	0.77			
	BT7	0.83			
	BT8	0.8			
	BT9	0.76			
Values	V 1	0.73	0.82	0.86	0.51
	V 2	0.67			
	V 3	0.73			
	V 4	0.73			
	V 5	0.73			
	V 6	0.63			
	V 7	0.68			

Based on traditional methods for assessing discriminant validity, an indicator's outer loadings on its associated construct must be greater than its cross-loadings with other constructs (Chin, 1998; Grégoire & Fisher, 2006; Henseler et al., 2009). Additionally, the Fornell-Larcker criterion states that the square root of the AVE of each construct should surpass its highest correlation with any other construct (Fornell & Larcker, 1981; Henseler et al., 2009). An HTMT value exceeding 0.90 (or > 0.85 for a conservative threshold) indicates a lack of discriminant validity (Henseler et al., 2009). The cross-loadings for all indicators meet the criteria, with each indicator loading higher on its intended construct than on any other construct, thus supporting the discriminant validity of the measurement model (see Annex D). In this research, there are slight deviations in the Fornell-Larcker criterion and HTMT ratio. Specifically, the square root of AVE for the GT construct is 0.85, which is slightly higher than the 0.76 for the BT construct (Table 3). Additionally, the HTMT value between GT and BT is 0.929, which is slightly above the traditional threshold of 0.9. However, when evaluating discriminant validity, slight deviations in the Fornell-Larcker criterion and HTMT ratio can be considered acceptable under certain circumstances. According to Voorhees et al. (2016), minor violations of the Fornell-Larcker criterion do not necessarily invalidate discriminant validity if other indicators, such as cross-loadings and theoretical justification, support the construct's distinctiveness. Similarly, Henseler et al. (2015) argue that an HTMT value slightly above the recommended threshold (e.g., 0.9) does not automatically imply a lack of discriminant validity, especially if the constructs are conceptually distinct and other validity measures are satisfactory. In this research, despite the GT construct having a slightly higher square root of AVE and the HTMT value between GT and BT being slightly above 0.850, the overall evidence supports discriminant validity.

Composite measurement model evaluation: To further investigate whether the model allows for the examination of relationships between the constructs, this research uses variance inflation factors (VIFs) to identify multicollinearity among the indicators and thus assess the degree of reliability of the model (Henseler et al., 2009). A VIF value below 10 is generally considered acceptable (J. F. Hair et al., 2010), with more conservative thresholds setting the maximum acceptable value below 5 (Kock & Lynn, 2012). In this model, all VIF values are below 5, ranging from 1.42 to 2.89 (see Annex. F). These values indicate minimal concern for potential multicollinearity.

Table 5 *- Fornell–Larcker criterion analysis and HTMT ratios*

	BT	**GT**	**MR**	**Values**
BT	**0.79**			
GT	0.85 (.93)	**0.77**		
MR	0.74 (.81)	0.72 (.79)	**0.76**	
Values	0.31 (.33)	0.31 (.34)	0.29 (.32)	**0.72**

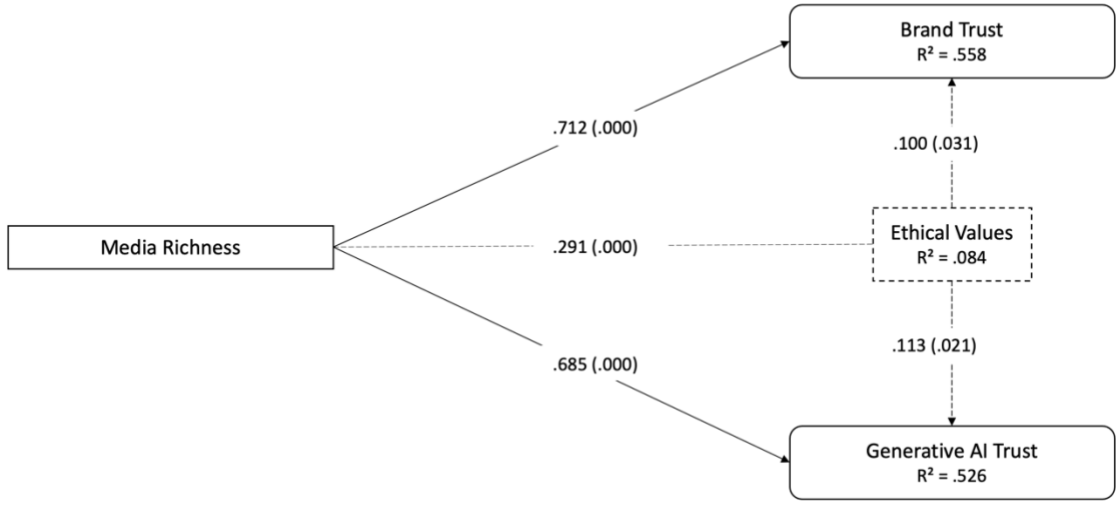
Notes: HTMT ratios are given in parentheses. The diagonal elements, highlighted in bold, represent the square roots of the average variance extracted (AVE) for each construct and its measures.

4.4.2 Inner Model results

An examination of the structural model fit indicates that the proposed model fits the data reasonably well (see Table 6). Henseler et al. (2015) introduced the SRMR as a fit index in PLS-SEM to prevent model misspecification. A value below 0.10, or ideally 0.08 according to J. F. Hair et al. (2010) (in a more conservative interpretation; refer to Hu & Bentler (1999)), indicates a good fit, as demonstrated in this study with an SRMR of 0.06. Addiotionally, with NFI values above 0.9 indicating an acceptable fit, the model achieving NFI = 0.84 meets the criteria for a model fit (J. F. Hair et al., 2010; Lohmöller, 1989).

The assessments of the structural model review the R^2 estimates, Stone-Geisser's Q2 value, effect size (f^2), path coefficients (β), and p-values, all of which are detailed in Figure 7 and Table 6.

Figure 7 *- Research Model with PLS-algorithm and bootstrapping results*



Note: The values correspond to the path coefficients. P-values are in parentheses.

Table 6 *- Structural Model Results*

Hypothesized relationship	**Proposed effect**	**Path coefficient**	**f^2**	**Results**
MR -> BT	Positive	.712 ***	1.050	H1: Supported
MR -> GT	Positive	.685 ***	.906	H2: Supported
MR -> Values	Positive	.291**	.092	H3: Supported
Values -> BT	Positive	.100*	.021	H4: Supported
Values -> GT	Positive	.113*	.024	H5: Supported
Variance explained	BT (R^2 = .558), GT (R^2 = .526), and Values (R^2 = .084)			

*Note: ***p < .001 **p<.01 *p<.05*

The R^2 values represent the proportion of variance in each variable explained by the model's predictors. Specifically, the model accounts for 4.84% (R^2 = .084) of the variance in Values, indicating weak predictive power, and 55.8% in BT (R^2 = .558) and 52.6% and GT (R^2 = .526) respectively, suggesting moderate predictive capability for these variables (J. F. Hair et al., 2011; Henseler et al., 2009; Sarstedt et al., 2014).

The effect sizes reveal significant relationships: MR exerts a large effect on BT (f^2 = 1.050) and GT (f^2 = .906), indicating substantial influence. It demonstrates a medium effect on Values (f^2 = .092), signifying moderate impact. In contrast, Values exhibit small effects on both BT (f^2 = .021) and GT (f^2 = .024), underscoring minor relationships. Results of 0.02, 0.15, and 0.35 correspond to small, medium, and large effect sizes (Cohen, 2013).

To evaluate the predictive relevance of the model, the cross-validated predictive ability test (CVPAT) is calculated with 10 folds and 10 repetitions for the overall model to assess out-of-sample predictive power (Liengaard et al., 2021; Sharma et al., 2023; Shmueli et al., 2019). Recent research indicates that the previously used blindfolding procedure, which utilized Stone-Geisser's Q^2 value (Geisser, 1974; Stone, 1974) as a criterion for evaluating the cross-validated predictive relevance of the PLS path model, does not adequately assess out-of-sample predictive power (J. T. Hair et al., 2022). The model exhibited a significantly lower average loss (i.e., higher predictive accuracy) compared to the naïve IA benchmark, thus meeting the minimum criterion for predictive validity (see Table 7) (Sharma et al., 2023). However, the model did not show a significantly lower average loss compared to the LM benchmark, indicating that it does not exhibit strong predictive validity (Sharma et al., 2023; Shmueli et al., 2019). To further evaluate the model's predictive performance, $Q^2_{predict}$ values are examined to compare prediction errors against simple mean predictions. The positive Q^2 values for all dependent variables further confirm the model's predictive validity (see Annex G) (Shmueli et al., 2019).

Table 7 - CVPAT Benchmark results for predictive model assessment

Focus on overall model	**PLS-SEM loss**	**Benchmark loss**	**Difference**	**p-value**
CVPATbenchmark_IA construct	.583	.785	-.175	< 0.05
CVPATbenchmark_LM construct	.583	.580	.003	.688 nsig

Note: loss = average loss

The path analysis results robustly support the proposed hypotheses with $p < \alpha$ ($\alpha = 0.05$) (J. F. Hair et al., 2010). MR significantly impacts BT ($\beta = .712$, $p < .001$) and GT ($\beta = .685$, p

< 0.001), indicating that MR plays a crucial role in enhancing these outcomes. Additionally, MR exhibits a moderate positive effect on Values ($\beta = .291$, $p < 0.01$), suggesting a meaningful relationship. The influence of Values on other constructs is also confirmed, with Values positively affecting BT ($\beta = .100$, $p < 0.05$) and GT ($\beta = .113$, $p < 0.05$), although these effects are weaker than those of MR. These findings confirm that MR is a significant predictor of BT and GT and positively influences Values. While Values also positively impacts BT and GT, its effect is less pronounced. Overall, the results affirm all hypotheses and underscore the influence of MR in shaping the dynamics among the constructs.

4.4.3 Mediation Analysis

A mediation analysis was performed to investigate the effects of Media Richness (MR) on Brand Trust (BT) and Generative AI Trust (GT), with the construct Values acting as the mediator, using the methodology outlined by Cepeda-Carrion et al. (2018). A bootstrapping procedure was utilized to calculate 97.5% confidence intervals for the indirect effects. Full mediation is indicated when only the indirect effect is significant, showing that the effect is fully transmitted through the mediator. Partial mediation involves both effects being significant, with complementary (same direction) or competitive (opposite directions) subtypes. No mediation is indicated if the indirect effect is not significant, which can be an only direct effect (significant direct path) or no effect (neither path significant)(Cepeda-Carrion & Nitzl, 2018).

Table 8 *- Mediation Analysis Results*

Effect	**Indirect effect**	**CI Indirect**		**VAF**	**Result**
		2.5%	97.5%		
(1) MR --> Values -> BT	.031^{nsig}	.004	.064	(4.17%)	No Mediation
(2) MR --> Values -> GT	.034*	.005	.066	4.72%	Complementary Partial Mediation

*Note: * indicates p-values less than 0.01 | VAF: variance accounted*

The direct effect of MR on BT was significant, with a coefficient (β) of 0.712 (**p < 0.001**), indicating a strong positive relationship between Media Richness and Brand Trust. Similarly, the direct effect of MR on GT was also significant, with a coefficient (β) of 0.686 (**p < 0.001**), demonstrating a robust positive relationship between MR and GT.

The indirect effect of MR on BT through Values was not significant, with a coefficient (β) of 0.031 and a p-value of 0.065, which exceeds the commonly accepted significance threshold. The confidence interval for this indirect effect ranged from 0.004 to 0.064, indicating that the true indirect effect could be very close to zero. Conversely, the indirect effect of MR on GT through Values was significant, with a coefficient (β) of 0.034 and a p-value of 0.039 (*p < 0.01*). The confidence interval for this indirect effect ranged from 0.005 to 0.066, indicating a small but significant mediation effect.

The total effect of MR on BT was significant, with a coefficient (β) of 0.743 (**p < 0.001**). This total effect is the sum of the direct and indirect effects, suggesting that MR has a substantial overall impact on BT. Similarly, the total effect of MR on GT was significant, with a coefficient (β) of 0.72 (**p < 0.001**), highlighting the overall influence of MR on GT.

To further understand the mediation effect, the Variance Accounted For (VAF) values were calculated. VAF indicates the proportion of the total effect that is mediated by the mediator variable. Approximately 4.72% of the total effect of MR on GT is mediated by Values. A full mediation can be assumed when the VAF value exceeds 80% (Cepeda-Carrion & Nitzl, 2018). The analysis therefore did not indicate full mediation in the relationship between MR and BT, as the direct effect remained significant and the indirect effect was not significant. Partial mediation was observed in the relationship between MR and GT. The significant indirect effect of MR on GT through Values (β = 0.034, *p < 0.01*) alongside the significant direct effect (β = 0.686, **p < 0.001**) suggests complementary partial mediation. This indicates that Values partially mediates the relationship between MR and GT, with both direct and indirect paths contributing to the overall effect. These findings are partly consistent with the mediation model assumptions, indicating that while Values serves as a mediator for the effect of MR on GT, it does not fully mediate the relationship between MR and BT. The results suggest that Media Richness impacts Brand Trust directly, whereas its influence on Generative AI Trust is partially mediated by Values.

5 Discussion

5.1 Key Findings

The primary aim of this study was to investigate how much information end-users need in order to build trust in generative AI and the organizations behind it, and what role their value systems play in this process. Table 8 represents a summary of results for hypothesis testing. Accordingly, the impact of Media Richness on end-user trust levels , specifically Brand Trust and Generative AI Trust, within the scope of generative AI was examined, with a particular focus on the mediating role of values. In doing so, this study takes an innovative and novel approach to enrich previous research. Drawing the bridge between what and how, this study not only aimed to explore whether the AI ethics principles established by the EU hold significant importance to individuals, but also to examine how much information end-users need in order to build trust in the technology and the organizations providing it. The focus is thus on testing a theoretical framework against actual consumer perceptions, therefore aiming to provide an experimental study to observe GAI interactions beyond the well established TAM model (Kelly, 2024; Ofosu-Ampong, 2024). This human-centered perspective is integrated by not only testing trust perceptions but also examining the value systems of the sample and their relevance in trust-building mechanisms (Ahmad et al., 2023), thus exploring the MRT beyond well-studied traditional communication contexts within the realm of GAI (Choung et al., 2023). This focus allows for a more profound understanding of the effectiveness of communication efforts and the underlying processes that shape perceptions, thereby enhancing overall comprehension.

Figure 8 *- Summary of results for hypothesis testing*

Hypothesis	Relationship	Result
H1	A direct relationship exists between Media Richness and Brand Trust	Accepted
H2	A direct relationship exists between Media Richness and Brand Trust	Accepted
H3	A direct relationship exists between the degree of Media Richness and values.	Accepted
H4	A direct relationship exists between values and Brand Trust.	Accepted
H5	A direct relationship exists between values and GAI trust.	Accepted
H6	Values mediate the relationship between Media Richness and Brand Trust.	Rejected
H7	Values mediate the relationship between Media Richness and Generative AI Trust.	Accepted

5.1.1 Degree of Media Richness on Trust Perceptions

RQ1: *How much information do end-users need in order to build trust in GAI and the organizations behind it?* The degree of Media Richness plays a noteworthy role in shaping end-users' trust in generative AI and the organizations behind it. The results indicate a positive and significant direct relationship between Media Richness and Brand Trust (H1), suggesting that richer media formats enhance Brand Trust. This finding aligns with previous research by Daft and Lengel (1986), who proposed that Media Richness can reduce uncertainty and ambiguity, thus fostering trust. Similarly, the study demonstrates a significant positive effect between Media Richness and GAI trust (H2). This supports the hypothesis that more immersive and detailed media formats lead to higher trust levels in generative AI systems, supporting findings by Mayer, Davis, and Schoorman (1995) on the importance of communication quality in trust formation.

Interestingly, all test groups showed positive trust perceptions for both the brand and the technology. In the first test group, which received a Company Statement with low Media Richness, only the company's measures for safe GAI implementation were communicated, without explanations of why these measures are important or in-depth technical details. The communication primarily conveyed that the company undertook actions to ensure the trustworthiness of generative AI, but it lacked detailed context regarding the why and how behind these actions. This suggests that even minimal information provided by a company about

its handling of GAI technology can be trust-building and does Media Richnessnot lead to user mistrust, avoiding negative effects similar to 'greenwashing' (Reinhardt, 2023; Wright, 2024; Wacksman, 2021). However, the level of trust increases with the provision of more information, as observed in test groups with moderate and high Media Richness. Notably, even the explicit technical explanation provided to the high Media Richness group led to higher trust levels, indicating that detailed information likely did not cause mistrust through user overload, contrary to previous concerns in the research. (Reinhardt, 2023; Wright, 2024; Wacksman, 2021). The results indicate that individuals appear to seek information about values, which subsequently fosters trust. Consequently, while Media Richness does play a significant role, the critical factor for generating trust is the provision of information about values. Regarding RQ1, it can be summarized that a high degree of Media Richness in a Company Statement, presented through a combination of communicating *the what* (the company's measures for trustworthy GAI implementation), *the why* (explanations of why these measures are important), and *the how* (detailed technical explanations of GAI) of their values, is most effective in building end-user trust in AI and the organization behind it.

5.1.2 Value Systems of End-Users

RQ2: *What values are important to people when it comes to trustworthy GAI design?*

In examining the values that users find crucial for trustworthy generative AI (GAI) design, several key patterns emerge. Privacy and Data Governance, Technical Robustness and Safety, and Transparency are prominently highlighted as priorities. This aligns with existing research that emphasizes these aspects as central to building trust in AI systems (Afroogh et al., 2024; Hleg AI, 2019; B. Li et al., 2023): Users show a strong preference for ensuring that AI systems are safe, well-regulated, and transparent. Privacy and Data Governance, along with Technical Robustness and Safety, are particularly significant for users, reflecting their concerns about security and adherence to privacy standards. Transparency is also highly valued, pointing to the need to reduce opacity in AI systems. While Accountability and Societal and Environmental Well-being are recognized as important, they are generally not prioritized as highly by users in their evaluations of AI systems. This suggests that these values, although acknowledged, are

seen as less critical on their own, rather than in relation to other concerns like security and transparency.

The agreement among participants from the three test groups regarding these values is noticeable, with a general consensus on their importance. This consistency underscores the need for generative AI systems to address these critical values to build and maintain end-user trust. Regarding RQ2, it can be therefore summarized that participants have a strong shared perspective and generally regard all presented values as important: Accountability, Transparency, Privacy and Data Governance, Diversity, Non-discrimination, and Fairness, Technical Robustness and Safety, Human Agency and Oversight, and Societal and Environmental Well-being.

5.1.3 Role of Values in building Trust

RQ3: What role do values play in the communicative process of building trust in generative AI? Values play a complex role in the communicative process of building trust in generative AI, particularly in how information about these values is conveyed through Media Richness. The analysis reveals that providing more information about values significantly strengthens the alignment between these values and the trust perceptions of consumers. This effect is particularly pronounced in the context of Generative AI Trust, where individuals who prioritize the values presented in the communication exhibit a stronger trust response when exposed to more detailed information about these values. This suggests that the effect of Media Richness on GAI Trust is amplified for those who find these values important, whereas for those who do not, the effect of MR is minimal. However, this amplifier effect does not hold in the context of Brand Trust. While MR influences the perception of values, and values, in turn, impact trust, the relationship between MR and Brand Trust does not differ significantly between individuals who place high or low importance on these values. Therefore, strategically emphasizing values such as accountability, transparency, and privacy can significantly enhance communication strategies around generative AI, particularly in fostering trust in the technology itself, where the depth of value-based communication resonates more profoundly with those who hold these values in high regard.

5.2 Implications

The study's findings offer valuable insights into how the depth of value-related information influences trust in generative AI systems and the organizations behind them, leading to the following theoretical and practical applications for effective communication strategies.

Theoretical Implications: The findings of this study provide several theoretical contributions to the field of trust in generative AI. First, while the study observed a positive relationship between Media Richness and both Brand Trust and GAI trust (H1 and H2), it extends the Media Richness Theory (Daft & Lengel, 1986) by focusing on the depth and immersiveness of information regarding values within a single media format (company statements), suggesting that such depth fosters trust. Second, the study highlights the mediating role of values in the relationship between the depth of information and trust (H7). This aligns with and expands on existing literature (Li, 2022; Morgan & Hunt, 1994), emphasizing the importance of value alignment in trust-building processes. By demonstrating that values significantly influence how the depth of information impacts trust perceptions, this research contributes to a deeper understanding of the mechanisms through which communication affects trust in AI. Third, the study integrates a human-centered perspective by examining the relevance of users' value systems in trust-building mechanisms, thereby enriching the traditional Technology Acceptance Model (TAM) and Media Richness Theory (MRT) frameworks (Ahmad et al., 2023; Choung et al., 2023). This approach provides a more holistic view of trust formation in generative AI, incorporating both technological and human factors. Overall, the theoretical implications emphasize the necessity of considering the depth of information about values in models of trust formation around generative AI, offering a nuanced perspective that can inform future research on trust in emerging technologies.

Practical Implications: The findings of this study offer practical insights for managers and organizations seeking to build trust in generative AI. First, the impact of detailed information about values suggests that organizations should invest in providing rich, comprehensive value-related content in their communications around GAI. Providing detailed, immersive information about the what, why, and how of GAI value implementation can enhance trust in both the technology and the organization. Second, the importance of value alignment indicates

that organizations should prioritize communicating their commitment to core values such as accountability, transparency, privacy, and data governance. Ensuring that these values are prominently featured in communication efforts can foster a deeper connection with end-users, as they perceive a strong alignment between their values and those of the organization. Third, the study suggests that while even minimal information about a company's measures for safe GAI implementation can build trust, more comprehensive information leads to higher trust levels. Organizations should strive to provide as much relevant information as possible without overwhelming users. Lastly, organizations should be aware that while values such as accountability and societal and environmental well-being are important, they may not be as central to users' trust evaluations as privacy, data governance, and transparency. Therefore, communication strategies should prioritize these key concerns while still addressing broader ethical considerations. By implementing these strategies, managers can enhance trust in their AI systems and the organization providing the technology, which can ultimately lead to more positive user experiences and stronger, more resilient relationships with their stakeholders.

5.3 Limitations & Future Research

Despite the significant insights gained from this study on the impact of Media Richness on end-user trust levels in generative AI, several limitations must be acknowledged, which also highlight opportunities for future research.

Sample Size and Diversity: The sample size and demographic diversity do not fully represent the broader population. The findings may be limited to the specific characteristics of the participants, impacting the generalizability of the results. Additionally, there was a limitation related to the average variance extracted (AVE) for the BT construct in the pre-study sample. This issue indicates that the construct may not fully capture the variance in its indicators, which could affect the validity and robustness of the study's findings. However, this limitation is somewhat mitigated by the small sample size of the pre-study, which may have influenced the AVE results. Future research should aim to include larger and more diverse samples to enhance generalizability and provide a more comprehensive view of trust in generative AI, considering various demographic, cultural, and professional backgrounds (Schoorman et al., 2007).

Self-Reported Data: The reliance on self-reported data for assessing trust and values may introduce bias. Participants' responses might be influenced by social desirability or their perceptions of what is expected, potentially affecting data accuracy. To mitigate this bias, it is crucial to conduct actual field studies or examine behavioral variables rather than relying solely on self-reported trust. For example, one could investigate whether individuals actually purchase more or engage more in certain behaviors, rather than just reporting their intentions or beliefs. Employing multiple methods, including qualitative approaches such as interviews or case studies, could offer a richer understanding of these constructs and help address potential biases.

Contextual Limitations: The research primarily focuses on generative AI in a specific context. Variations in AI applications across different sectors or cultural settings might yield different results, limiting the study's applicability to other scenarios. Future research should explore the impact of Media Richness on trust across various contexts and applications, such as healthcare, finance, and education, to provide industry-specific insights and recommendations.

Temporal Constraints: The study captures a snapshot of perceptions at a particular point in time. As generative AI technology evolves and societal attitudes change, the findings may become outdated. Longitudinal studies could provide a more comprehensive understanding of how trust dynamics evolve over time and with technological advancements (Ployhart & Vandenberg, 2010).

Measurement Limitations: The constructs of Media Richness, values, and trust were measured using specific instruments that may not capture all nuances of these complex constructs. It is important to note that the stimuli used were self-created, and the validation of these operationalizations should be tested in further studies. Additionally, the relationship between Media Richness and values was explored, which means that general statements about MR alone might not be possible. Future research should test the mediation effects using stimuli that do not include the dependent variables in question. Investigating emerging media formats and their effects on trust, such as videos, infographics, and interactive content, could offer valuable insights into the effectiveness of conveying complex information.

Exploring Additional Mediators and Moderators: Future studies could explore not only additional mediators but also moderators that influence how Media Richness affects trust. While perceived risk, user experience, and emotional responses are potential mediators that

could provide a more nuanced understanding of the mechanisms through which Media Richness influences trust (McKnight et al., 2002a), it is equally important to consider various moderators. Moderators such as individual personality traits, demographic characteristics (e.g., age, education), and general attributes like uncertainty or trust in people could significantly impact the relationship between Media Richness and trust. These factors may alter how Media Richness is perceived and how it ultimately affects trust, offering a more comprehensive view of the dynamics at play, leading to a more detailed and accurate understanding of these interactions.

Incentive-Driven Participation Bias: Incentives were employed to enhance participant engagement and increase response rates. However, this strategy may introduce a limitation, as it could result in participants taking part primarily for the rewards rather than out of intrinsic interest in the study.

Bibliographic References

Abillama, N., Mills, S., Boison, G., & Carrasco, M. (2021, July 21). *Unlocking the Value of AI-Powered Government*. Boston Consulting. https://www.bcg.com/publications/2021/unlocking-value-ai-in-government

Adobe. (2024). *Adobe's Commitment to AI Ethics*. https://www.adobe.com/content/dam/cc/en/ai-ethics/pdfs/Adobe-AI-Ethics-Principles.pdf

Adomako, S., & Nguyen, N. P. (2023). Green creativity, responsible innovation, and product innovation performance: A study of entrepreneurial firms in an emerging economy. *Business Strategy and the Environment*, *32*(7), 4413–4425. https://doi.org/10.1002/bse.3373

Afroogh, S., Akbari, A., Malone, E., Kargar, M., & Alambeigi, H. (2024). Trust in AI: Progress, Challenges, and Future Directions. *ArXiv*, *abs/2403.14680*. https://doi.org/https://doi.org/10.48550/arXiv.2403.14680

Agarwal, P. K. (2018). Public Administration Challenges in the World of AI and Bots. *Public Administration Review*, *78*(6), 917–921. https://doi.org/10.1111/puar.12979

Ågerfalk, P. J. (2020). Artificial intelligence as digital agency. *European Journal of Information Systems*, *29*(1), 1–8. https://doi.org/10.1080/0960085X.2020.1721947

Ahmad, K., Abdelrazek, M., Arora, C., Bano, M., & Grundy, J. (2023). Requirements engineering for artificial intelligence systems: A systematic mapping study. *Information and Software Technology*, *158*, 107176. https://doi.org/10.1016/j.infsof.2023.107176

Ahmed, T., Karmaker, C. L., Nasir, S. B., Moktadir, Md. A., & Paul, S. K. (2023). Modeling the artificial intelligence-based imperatives of industry 5.0 towards resilient supply chains: A post-COVID-19 pandemic perspective. *Computers & Industrial Engineering*, *177*, 109055. https://doi.org/10.1016/j.cie.2023.109055

Alarcon, G. M., Lyons, J. B., Christensen, J. C., Klosterman, S. L., Bowers, M. A., Ryan, T. J., Jessup, S. A., & Wynne, K. T. (2018). The effect of propensity to trust and perceptions of trustworthiness on trust behaviors in dyads. *Behavior Research Methods*, *50*(5), 1906–1920. https://doi.org/10.3758/s13428-017-0959-6

Al-Khaldi, M. A., & Olusegun Wallace, R. S. (1999). The influence of attitudes on personal computer utilization among knowledge workers: the case of Saudi Arabia. *Information & Management*, *36*(4), 185–204. https://doi.org/10.1016/S0378-7206(99)00017-8

Allen, C., & Wallach, W. (2012). Moral machines: Contradiction in terms of abdication of human responsibility? . In P. Lin, K. Abney, & G. Bekey (Eds.), *Robot ethics: the ethical and social implications of robotics* (First MIT Press, pp. 55–68). The MIT Press.

Ameen, N., Tarhini, A., Reppel, A., & Anand, A. (2021). Customer experiences in the age of artificial intelligence. *Computers in Human Behavior*, *114*, 106548. https://doi.org/10.1016/j.chb.2020.106548

Araujo, T., Helberger, N., Kruikemeier, S., & de Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & SOCIETY*, *35*(3), 611–623. https://doi.org/10.1007/s00146-019-00931-w

Arogyaswamy, B. (2020). Big tech and societal sustainability: an ethical framework. *AI & SOCIETY*, *35*(4), 829–840. https://doi.org/10.1007/s00146-020-00956-6

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for Artificial Intelligence and Digital technologies. *International Journal of Information Management*, *62*, 102433. https://doi.org/10.1016/j.ijinfomgt.2021.102433

Ashok, M., Narula, R., & Martinez-Noya, A. (2016). How do collaboration and investments in knowledge management affect process innovation in services? *Journal of Knowledge Management*, *20*(5), 1004–1024. https://doi.org/10.1108/JKM-11-2015-0429

Ashoori, M., & Weisz, J. D. (2019). *In AI We Trust? Factors That Influence Trustworthiness of AI-infused Decision-Making Processes*.

Attard-Frost, B., De los Ríos, A., & Walters, D. R. (2023). The ethics of AI business practices: a review of 47 AI ethics guidelines. *AI and Ethics*, *3*(2), 389–406. https://doi.org/10.1007/s43681-022-00156-6

Bach, T. A., Khan, A., Hallock, H., Beltrão, G., & Sousa, S. (2024). A Systematic Literature Review of User Trust in AI-Enabled Systems: An HCI Perspective. *International Journal of Human–Computer Interaction*, *40*(5), 1251–1266. https://doi.org/10.1080/10447318.2022.2138826

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74–94. https://doi.org/10.1007/BF02723327

Bandi, A., Adapa, P. V. S. R., & Kuchi, Y. E. V. P. K. (2023). The Power of Generative AI: A Review of Requirements, Models, Input–Output Formats, Evaluation Metrics, and Challenges. *Future Internet*, *15*(8), 260. https://doi.org/10.3390/fi15080260

Bartoletti, I. (2020). *An Artificial Revolution: On Power, Politics and AI.* . The Indigo Press.

Bathaee, Y. (2018). The artificial intelligence black box and the failure of intent and causation. *Harvard Journal of Law & Technology*, *31*(2), 890–938.

Bauer, P. (2019). *Conceptualizing Trust and Trustworthiness* (61; Revised Version of Working Paper).

Becker, D. (2018, April). Possibilities to Improve Online Mental Health Treatment: Recommendations for Future Research and Developments. *Future of Information and Communication Conference (FICC)*.

Bedué, P., & Fritzsche, A. (2022). Can we trust AI? An empirical investigation of trust requirements and guide to successful AI adoption. *Journal of Enterprise Information Management*, *35*(2), 530–549. https://doi.org/10.1108/JEIM-06-2020-0233

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Special Issue Editor's Comments: Managing Artificial Intelligence. *Management Information Systems Quarterly*, *45*(3), 1433–1450.

Berg, J. M., Raj, M., & Seamans, R. (2023a). Capturing Value from Artificial Intelligence. *Academy of Management Discoveries*, *9*(4), 424–428. https://doi.org/10.5465/amd.2023.0106

Berg, J. M., Raj, M., & Seamans, R. (2023b). Capturing Value from Artificial Intelligence. *Academy of Management Discoveries*, *9*(4), 424–428. https://doi.org/10.5465/amd.2023.0106

Berglind, N., Fadia, A., & Isherwood, T. (2022, July 25). *The potential value of AI—and how governments could look to capture it*. McKinsey. https://www.mckinsey.com/industries/public-sector/our-insights/the-potential-value-of-ai-and-how-governments-could-look-to-capture-it

Blau, P. M. (2017). *Exchange and Power in Social Life*. Routledge. https://doi.org/10.4324/9780203792643

Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L. D., Monfort, M., Muller, U., Zhang, J., Zhang, X., Zhao, J., & Zieba, K. (2016). *End to End Learning for Self-Driving Cars*.

Boscardin, C. K., Gin, B., Golde, P. B., & Hauer, K. E. (2023). ChatGPT and Generative Artificial Intelligence for Medical Education: Potential Impact and Opportunity. *Academic Medicine*. https://doi.org/10.1097/ACM.0000000000005439

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies.* . Oxford University Press.

Botha, J., & Pieterse, H. (2020, March). Fake News and Deepfakes: A Dangerous Threat for 21st Century Information Security. *15t International Conference on Cyber Warfare and Security*.

Brynjolfsson, E., Li, D., & Raymond, L. (2023). *Generative AI at Work*. https://doi.org/10.3386/w31161

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies.* . WW Norton & Company.

Bughin, J., Seong, J., Manyika, J., Chui, M. , & Joshi, R. (2018). *Notes from the AI frontier: Modeling the impact of AI on the world economy*. https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy

Buolamwini, J. A. (2017). *Gender shades : intersectional phenotypic and demographic evaluation of face datasets and gender classifiers*. Massachusetts Institute of Technology.

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *FAT*.

Calhoun, C. S., Bobko, P., Gallimore, J. J., & Lyons, J. B. (2019). Linking precursors of interpersonal trust to human-automation trust: An expanded typology and exploratory experiment. *Journal of Trust Research*, *9*(1), 28–46. https://doi.org/10.1080/21515581.2019.1579730

Callegaro, M., Manfreda, K. L., & Vehovar, V. (2015). *Web survey methodology*. Sage.

Carlson, J. R., & Zmud, R. W. (1999). CHANNEL EXPANSION THEORY AND THE EXPERIENTIAL NATURE OF MEDIA RICHNESS PERCEPTIONS. *Academy of Management Journal*, *42*(2), 153–170. https://doi.org/10.2307/257090

Carter, S. P., & Rogers, E. S. (2008). A review of government trust and transparency in public institutions. *Public Administration Review*, *68*(1), 39–47.

Carvalho, A., Levitt, A., Levitt, S., Khaddam, E., & Benamati, J. (2019). Off-The-Shelf Artificial Intelligence Technologies for Sentiment and Emotion Analysis: A Tutorial on Using IBM Natural Language Processing. *Communications of the Association for Information Systems*, 918–943. https://doi.org/10.17705/1CAIS.04443

Castelli, M., & Manzoni, L. (2022). Special Issue: Generative Models in Artificial Intelligence and Their Applications. *Applied Sciences*, *12*(9), 4127. https://doi.org/10.3390/app12094127

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, *56*(5), 809–825. https://doi.org/10.1177/0022243719851788

Cepeda-Carrion, G., & Nitzl, C. (2018). Mediation Analyses in Partial Least Squares Structural Equation Modeling. Guidelines and Empirical Examples. In H. Latan, R. Noonan, & J. Roldán (Eds.), *Partial Least Squares Structural Equation Modeling: Basic Concepts, Methodological Issues and Applications.* Springer.

Chang, S. E., Liu, A. Y., & Shen, W. C. (2017). User trust in social networking services: A comparison of Facebook and LinkedIn. *Computers in Human Behavior*, *69*, 207–217. https://doi.org/10.1016/j.chb.2016.12.013

Chaudhuri, A., & Holbrook, M. B. (2001). The Chain of Effects from Brand Trust and Brand Affect to Brand Performance: The Role of Brand Loyalty. *Journal of Marketing*, *65*(2), 81–93. https://doi.org/10.1509/jmkg.65.2.81.18255

Cheatham, B., Javanmardian, K., & Samandari, H. (2019). Confronting the Risks of Artificial Intelligence. *McKinsey Quarterly 2*, *38.* https://www.mckinsey.com/capabilities/quantumblack/our-insights/confronting-the-risks-of-artificial-intelligence

Cheng, J.-Z., Ni, D., Chou, Y.-H., Qin, J., Tiu, C.-M., Chang, Y.-C., Huang, C.-S., Shen, D., & Chen, C.-M. (2016). Computer-Aided Diagnosis with Deep Learning Architecture:

Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans. *Scientific Reports*, *6*(1), 24454. https://doi.org/10.1038/srep24454

Chi, O. H., Jia, S., Li, Y., & Gursoy, D. (2021). Developing a formative scale to measure consumers' trust toward interaction with artificially intelligent (AI) social robots in service delivery. *Computers in Human Behavior*, *118*, 106700. https://doi.org/10.1016/j.chb.2021.106700

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern Methods for Business Research*, *295*(2), 295–336.

Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence*, *6*, 100197. https://doi.org/10.1016/j.caeai.2023.100197

Cho, C. H., Phillips, J. R., Hageman, A. M., & Patten, D. M. (2009). Media richness, user trust, and perceptions of corporate social responsibility. *Accounting, Auditing & Accountability Journal*, *22*(6), 933–952. https://doi.org/10.1108/09513570910980481

Choung, H., David, P., & Ross, A. (2023a). Trust and ethics in AI. *AI & SOCIETY*, *38*(2), 733–745. https://doi.org/10.1007/s00146-022-01473-4

Choung, H., David, P., & Ross, A. (2023b). Trust and ethics in AI. *AI & SOCIETY*, *38*(2), 733–745. https://doi.org/10.1007/s00146-022-01473-4

Choung, H., David, P., & Ross, A. (2023c). Trust in AI and Its Role in the Acceptance of AI Technologies. *International Journal of Human–Computer Interaction*, *39*(9), 1727–1739. https://doi.org/10.1080/10447318.2022.2050543

Clayton, J. (2023, May 17). *Sam Altman: CEO of OpenAI calls for US to regulate artificial intelligence*. BBC.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Cruz-Cárdenas, J., Zabelina, E., Deyneka, O., Guadalupe-Lanas, J., & Velín-Fárez, M. (2019). Role of demographic factors, attitudes toward technology, and cultural values in the prediction of technology-based consumer behaviors: A study in developing and emerging countries. *Technological Forecasting and Social Change*, *149*, 119768. https://doi.org/10.1016/j.techfore.2019.119768

Daft, R. L., & Lengel, R. H. (1986). Organizational Information Requirements, Media Richness and Structural Design. *Management Science*, *32*(5), 554–571. https://doi.org/10.1287/mnsc.32.5.554

Dastin, J. (2022). Amazon scraps secret AI recruiting tool that showed bias against women. In *Ethics of data and analytics* (pp. 296–299). Auerbach Publications.

David, P., Choung, H., & Seberger, J. S. (2024). Who is responsible? US Public perceptions of AI governance through the lenses of trust and ethics. *Public Understanding of Science*. https://doi.org/10.1177/09636625231224592

De Mooij, M. (2019). *Global marketing and advertising: Understanding cultural paradoxes* (5th ed.). SAGE Publications.

Degas, A., Islam, M. R., Hurter, C., Barua, S., Rahman, H., Poudel, M., Ruscio, D., Ahmed, M. U., Begum, S., Rahman, M. A., Bonelli, S., Cartocci, G., Di Flumeri, G., Borghini, G., Babiloni, F., & Aricó, P. (2022). A Survey on Artificial Intelligence (AI) and eXplainable AI in Air Traffic Management: Current Trends and Development with Future Research Trajectory. *Applied Sciences*, *12*(3), 1295. https://doi.org/10.3390/app12031295

Delgado-Ballester, E., & Alemán, L. J. (2005). Does brand trust matter to brand equity? *Journal of Product & Brand Management*, *14*(3), 187–196. https://doi.org/10.1108/10610420510601058

Deloitte. (2022). *How to create an effective AI strategy* . Perspectives on AI. https://www2.deloitte.com/us/en/pages/technology/articles/effective-ai-strategy.html

Devitt, K. (2018). Trustworthiness of Autonomous Systems. In J. Abbass, D. Scholz, & J. Reid (Eds.), *Foundations of Trusted Autonomy* (1st ed., Vol. 117, pp. 161–184). Springer International Publishing. https://doi.org/10.1007/978-3-319-64816-3_9

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033

Ditsa, G., & MacGregor, G. E. (1995). Models of User Perceptions, Expectationsand Acceptance of Information Systems. *Sixth International Conference of TheInformation Resource Management Association*, 221–229.

Doney, P. M., & Cannon, J. P. (1997). An Examination of the Nature of Trust in Buyer-Seller Relationships. *Journal of Marketing*, *61*(2), 35. https://doi.org/10.2307/1251829

Doraiswamy, P. M., London, E., Varnum, P., Harvey, B., Saxena, S., Tottman, S., Campbell, S. P., Fernandez Ibanez, A., Manji, H., Al Olama, M. A. A. S., Chou, I.-H., Herrman, H., Jeong, S., Le, T., Montojo, C., Reve, B., Rommelfanger, K. S., Stix, C., Thakor, N., … Candeias, V. (2019). Empowering 8 Billion Minds: Enabling Better Mental Health for All via the Ethical Adoption of Technologies. *NAM Perspectives*. https://doi.org/10.31478/201910b

Dörner, D. (1980). On the difficulties people have in dealing with complexity. *Simulation & Games*, *11*(1), 87–106.

Eisingerich, A. B., & Bell, S. J. (2008). Perceived Service Quality and Customer Trust. *Journal of Service Research*, *10*(3), 256–268. https://doi.org/10.1177/1094670507310769

El Atillah, I. (2023, March 23). *Man ends his life after an AI chatbot "encouraged" him to sacrifice himself to stop climate change*. Euronews. https://www.euronews.com/next/2023/03/31/man-ends-his-life-after-an-ai-chatbot-encouraged-him-to-sacrifice-himself-to-stop-climate-

Fainman, A. A. (2019). The problem with Opaque AI. *The Thinker*, *82*(4), 44–55. https://doi.org/10.36615/thethinker.v82i4.373

Fast, E., & Horvitz, E. (2017). Long-Term Trends in the Public Perception of Artificial Intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*, *31*(1). https://doi.org/10.1609/aaai.v31i1.10635

Ferreira, D. B., Giovannini, C., Gromova, E., & da Rocha Schmidt, G. (2022). Arbitration chambers and trust in technology provider: Impacts of trust in technology intermediated dispute resolution proceedings. *Technology in Society*, *68*, 101872. https://doi.org/10.1016/j.techsoc.2022.101872

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, *66*(1), 111–126. https://doi.org/10.1007/s12599-023-00834-7

Floridi, L. (2007). A Look into the Future Impact of ICT on Our Lives. *The Information Society*, *23*(1), 59–64. https://doi.org/10.1080/01972240601059094

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Fornell, C., & Larcker, D. F. (1981). Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics. *Journal of Marketing Research*, *18*(3), 382. https://doi.org/10.2307/3150980

Fournier, S. (1998). Consumers and Their Brands: Developing Relationship Theory in Consumer Research. *Journal of Consumer Research*, *24*(4), 343–353. https://doi.org/10.1086/209515

French, A., Shim, J. P., Risius, M., R. Larsen, K., & Jain, H. (2021). The 4th Industrial Revolution Powered by the Integration of AI, Blockchain, and 5G. *Communications of the Association for Information Systems*, *49*(1), 266–286. https://doi.org/10.17705/1CAIS.04910

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2021). Will Humans-in-the-Loop Become Borgs? Merits and Pitfalls of Working with AI. *MIS Quarterly*, *45*(3), 1527–1556. https://doi.org/10.25300/MISQ/2021/16553

Fulk, J., Steinfield, C., Schmitz, J., & Power, G. (1987). A Social Information Processing Model of Media Use in Organizations. *Communication Research*, *14*(5), 529–552. https://doi.org/10.1177/009365087014005005

Ganster, D. C., Hennessey, H. W., & Luthans, F. (1983). Social Desirability Response Effects: Three Alternative Models. *Academy of Management Journal*, *26*(2), 321–331. https://doi.org/10.2307/255979

Gefen, Karahanna, & Straub. (2003). Trust and TAM in Online Shopping: An Integrated Model. *MIS Quarterly*, *27*(1), 51. https://doi.org/10.2307/30036519

Geisser, S. (1974). A Predictive Approach to the Random Effects Model. *Biometrika*, *61*(1), 101–107.

Gelderman, M. (1998). The relation between user satisfaction, usage of information systems and performance. *Information & Management*, *34*(1), 11–18. https://doi.org/10.1016/S0378-7206(98)00044-5

Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. *Computers in Human Behavior*, *115*, 106607. https://doi.org/10.1016/j.chb.2020.106607

Goldman Sachs. (2023, April 5). *Generative AI could raise global GDP by 7%*. https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html

Golembiewski, R. T., & McConkie, M. (1975). The centrality of interpersonal trust in group processes. *Theories of Group Processes* , 131–185.

Gordon, F. (2018). The Significance and Impact of the Media in Contemporary Society. In *Children, Young People and the Press in a Transitioning Society* (pp. 17–46). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-60682-2_2

Granulo, A., Fuchs, C., & Puntoni, S. (2021). Preference for Human (vs. Robotic) Labor is Stronger in Symbolic Consumption Contexts. *Journal of Consumer Psychology*, *31*(1), 72–80. https://doi.org/10.1002/jcpy.1181

Grégoire, Y., & Fisher, R. J. (2006). The effects of relationship quality on customer retaliation. Marketing Letters, 17, 31-46. *Marketing Letters*, *17*, 31–46.

Grewal, D., Roggeveen, A. L., & Nordfält, J. (2017). The Future of Retailing. *Journal of Retailing*, *93*(1), 1–6. https://doi.org/10.1016/j.jretai.2016.12.008

Guerreiro, J., Rita, P., & Trigueiros, D. (2015). Attention, emotions and cause-related marketing effectiveness. *European Journal of Marketing*, *49*(11/12), 1728–1750. https://doi.org/10.1108/EJM-09-2014-0543

Gulati, S., Sousa, S., & Lamas, D. (2019). Design, development and evaluation of a human-computer trust scale. *Behaviour & Information Technology*, *38*(10), 1004–1015. https://doi.org/10.1080/0144929X.2019.1656779

Gummer, T., Roßmann, J., & Silber, H. (2021). Using Instructed Response Items as Attention Checks in Web Surveys: Properties and Implementation. *Sociological Methods & Research*, *50*(1), 238–264. https://doi.org/10.1177/0049124118769083

Hadj, T. B. (2020). Effects of corporate social responsibility towards stakeholders and environmental management on responsible innovation and competitiveness. *Journal of Cleaner Production*, *250*, 119490. https://doi.org/10.1016/j.jclepro.2019.119490

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, *30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hair, J. F., Black, W. C., Barbin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis* (7th ed.). Pearson Education Limited.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152. https://doi.org/10.2753/MTP1069-6679190202

Hair, J. T., Hult, G. M., Ringle, C., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM).* (3rd ed.). Sage Publishing.

Haleem, A., Javaid, M., Asim Qadri, M., Pratap Singh, R., & Suman, R. (2022). Artificial intelligence (AI) applications for marketing: A literature-based study. *International Journal of Intelligent Networks*, *3*, 119–132. https://doi.org/10.1016/j.ijin.2022.08.005

Haupt, C. E., & Marks, M. (2023). AI-Generated Medical Advice—GPT and Beyond. *JAMA*, *329*(16), 1349. https://doi.org/10.1001/jama.2023.5321

Helberger, N., Araujo, T., & de Vreese, C. H. (2020). Who is the fairest of them all? Public attitudes and expectations regarding automated decision-making. *Computer Law & Security Review*, *39*, 105456. https://doi.org/10.1016/j.clsr.2020.105456

Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., van den Hoven, J., Zicari, R. V., & Zwitter, A. (2019). Will Democracy Survive Big Data and Artificial Intelligence? In *Towards Digital Enlightenment* (pp. 73–98). Springer International Publishing. https://doi.org/10.1007/978-3-319-90869-4_7

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: updated guidelines. *Industrial Management & Data Systems*, *116*(1), 2–20. https://doi.org/10.1108/IMDS-09-2015-0382

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In *Advances in International Marketing* (pp. 277–319). https://doi.org/10.1108/S1474-7979(2009)0000020014

Hibben, K. C., Felderer, B., & Conrad, F. G. (2022). Respondent commitment: applying techniques from face-to-face interviewing to online collection of employment data. *International Journal of Social Research Methodology*, *25*(1), 15–27. https://doi.org/10.1080/13645579.2020.1826647

Highhouse, S. (2008). Stubborn Reliance on Intuition and Subjectivity in Employee Selection. *Industrial and Organizational Psychology*, *1*(3), 333–342. https://doi.org/10.1111/j.1754-9434.2008.00058.x

Hleg AI. (2019). Ethics guidelines for trustworthy AI. In *European Commission*. https://www.europarl.europa.eu/cmsdata/196377/AI%20HLEG_Ethics%20Guidelines%20for%20Trustworthy%20AI.pdf

Hoff, K. A., & Bashir, M. (2015). Trust in Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Hoffmann, H., & Söllner, M. (2014). Incorporating behavioral trust theory into system development for ubiquitous applications. *Personal and Ubiquitous Computing*, *18*(1), 117–128. https://doi.org/10.1007/s00779-012-0631-1

Hollingshead, A. B., & McGrath, J. E. (1993). Putting the "group" back in group support systems: Some theoretical issues about dynamic processes in groups with technological enhancements. . *Group Support Systems: New Perspectives*, 78–96.

Hon, L. C., & Grunig, J. E. (1999). Guidelines for Measuring Relationships in Public Relations. The Institute for Public Relations. In *The Institute for Public Relations*. https://www.instituteforpr.org/wp-content/uploads/Guidelines_Measuring_Relationships.pdf

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, M.-H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, *21*(2), 155–172. https://doi.org/10.1177/1094670517752459

Huang, P. H., & Zhang, Y. S. (2020). Media richness and adoption intention of voice assistants: a cross-cultural study. *International Journal of Multinational Corporation Strategy*, *3*(2), 108. https://doi.org/10.1504/IJMCS.2020.114688

Iacobucci, D., & Churchill, G. A. (2010). *Marketing research: Methodological Foundations*. South Western Educational Publishing.

Ibáñez, J. C., & Olmeda, M. V. (2022). Operationalising AI ethics: how are companies bridging the gap between practice and principles? An exploratory study. *AI & SOCIETY*, *37*(4), 1663–1687. https://doi.org/10.1007/s00146-021-01267-0

Ishii, K., Lyons, M. M., & Carr, S. A. (2019). Revisiting media richness theory for today and future. *Human Behavior and Emerging Technologies*, *1*(2), 124–131. https://doi.org/10.1002/hbe2.138

Ismatullaev, U. V. U., & Kim, S.-H. (2024a). Review of the Factors Affecting Acceptance of AI-Infused Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *66*(1), 126–144. https://doi.org/10.1177/00187208211064707

Ismatullaev, U. V. U., & Kim, S.-H. (2024b). Review of the Factors Affecting Acceptance of AI-Infused Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *66*(1), 126–144. https://doi.org/10.1177/00187208211064707

Jacovi, A., Marasović, A., Miller, T., & Goldberg, Y. (2021a). Formalizing Trust in Artificial Intelligence. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 624–635. https://doi.org/10.1145/3442188.3445923

Jacovi, A., Marasović, A., Miller, T., & Goldberg, Y. (2021b). Formalizing Trust in Artificial Intelligence. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 624–635. https://doi.org/10.1145/3442188.3445923

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research. *Journal of Consumer Research*, *30*(2), 199–218. https://doi.org/10.1086/376806

Jo Hatch, M., & Schultz, M. (2003). Bringing the corporation into corporate branding. *European Journal of Marketing*, *37*(7/8), 1041–1064. https://doi.org/10.1108/03090560310477654

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Johnson-Eilola, J., & Selber, S. A. (2023). Technical Communication as Assemblage. *Technical Communication Quarterly*, *32*(1), 79–97. https://doi.org/10.1080/10572252.2022.2036815

Johnson-George, C., & & Swap, W. C. (1982). Measurement of specific interpersonal trust: construction and validation of a scale to assess trust in a specific other. *Journal of Personality and Social Psychology*, *43*, 1306–1317.

Jussupow, E., Benbasat, I., & Heinzl, A. (2020). WHY ARE WE AVERSE TOWARDS ALGORITHMS? A COMPREHENSIVE LITERATURE REVIEW ON ALGORITHM AVERSION. *ECIS 2020 Proceedings Research Papers*, *168*.

Kaplan, R. S., Norton, D. P., & Ansari, S. (2010). The Execution Premium: Linking Strategy to Operations for Competitive Advantage. *The Accounting Review*, *85*(4), 1475–1477. https://doi.org/10.2308/accr.2010.85.4.1475

Kar, A. K., Varsha, P. S., & Rajan, S. (2023). Unravelling the Impact of Generative Artificial Intelligence (GAI) in Industrial Applications: A Review of Scientific and Grey Literature. *Global Journal of Flexible Systems Management*, *24*(4), 659–689. https://doi.org/10.1007/s40171-023-00356-x

Kashive, N., Powale, L., & Kashive, K. (2020). Understanding user perception toward artificial intelligence (AI) enabled e-learning. *The International Journal of Information and Learning Technology*, *38*(1), 1–19. https://doi.org/10.1108/IJILT-05-2020-0090

Kaye, S.-A., Lewis, I., Forward, S., & Delhomme, P. (2020). A priori acceptance of highly automated cars in Australia, France, and Sweden: A theoretically-informed investigation guided by the TPB and UTAUT. *Accident Analysis & Prevention*, *137*, 105441. https://doi.org/10.1016/j.aap.2020.105441

Kayser-Bril, N. (2020). *Google apologizes after its Vision AI produced racist results.* AlgorithmWatch. https:// algorithmwatch.org/en/google-vision-racism/

Keding, C., & Meissner, P. (2021). Managerial overreliance on AI-augmented decision-making processes: How the use of AI-based advisory systems shapes choice behavior in R&D investment decisions. *Technological Forecasting and Social Change*, *171*, 120970. https://doi.org/10.1016/j.techfore.2021.120970

Kelly, S., Kaye, S.-A., & Oviedo-Trespalacios, O. (2023). What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telematics and Informatics*, *77*, 101925. https://doi.org/10.1016/j.tele.2022.101925

Khan, G. (2023, November 23). *Will AI-generated images create a new crisis for fact-checkers? Experts are not so sure*. Reuters Institute.

Kock, N. (2005). Media Richness or Media Naturalness? The Evolution of Our Biological Communication Apparatus and Its Influence on Our Behavior Toward E-Communication Tools. *IEEE Transactions on Professional Communication*, *48*(2), 117–130. https://doi.org/10.1109/TPC.2005.849649

Kock, N., & Lynn, G. (2012). Lateral Collinearity and Misleading Results in Variance-Based SEM: An Illustration and Recommendations. *Journal of the Association for Information Systems*, *13*(7), 546–580. https://doi.org/10.17705/1jais.00302

Kohli, P., & Chadha, A. (2020). *Enabling Pedestrian Safety Using Computer Vision Techniques: A Case Study of the 2018 Uber Inc. Self-driving Car Crash*. 261–279. https://doi.org/10.1007/978-3-030-12388-8_19

Kraus, S., Kumar, S., Lim, W. M., Kaur, J., Sharma, A., & Schiavone, F. (2023). From moon landing to metaverse: Tracing the evolution of Technological Forecasting and Social Change. *Technological Forecasting and Social Change*, *189*, 122381. https://doi.org/10.1016/j.techfore.2023.122381

Kumar, S., Lim, W. M., Pandey, N., & Christopher Westland, J. (2021). 20 years of Electronic Commerce Research. *Electronic Commerce Research*, *21*(1), 1–40. https://doi.org/10.1007/s10660-021-09464-1

Kumar, V., & Shah, D. (2004). Building and sustaining profitable customer loyalty for the 21st century. *Journal of Retailing*, *80*(4), 317–329. https://doi.org/10.1016/j.jretai.2004.10.007

Kung, F. Y. H., Kwok, N., & Brown, D. J. (2018). Are Attention Check Questions a Threat to Scale Validity? *Applied Psychology*, *67*(2), 264–283. https://doi.org/10.1111/apps.12108

Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy*, *44*(6), 101976. https://doi.org/10.1016/j.telpol.2020.101976

LaFollette, H. (1996). *Personal Relationship: Love, Identity, and Morality*. Blackwell Publishers.

Lakkaraju, H., & Bastani, O. (2020). "How do I fool you?" Manipulating User Trust via Misleading Black Box Explanations. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 79–85. https://doi.org/10.1145/3375627.3375833

Lancelot Miltgen, C., Popovič, A., & Oliveira, T. (2013). Determinants of end-user acceptance of biometrics: Integrating the "Big 3" of technology acceptance with privacy context. *Decision Support Systems, 56,* 103–114. https://doi.org/10.1016/j.dss.2013.05.010

Lankton, N., McKnight, D. H., & Tripp, J. (2015). Technology, Humanness, and Trust: Rethinking Trust in Technology. *Journal of the Association for Information Systems, 16*(10), 880–918. https://doi.org/10.17705/1jais.00411

Lanus, E., Freeman, L. J., Richard Kuhn, D., & Kacker, R. N. (2021). Combinatorial Testing Metrics for Machine Learning. *2021 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, 81–84. https://doi.org/10.1109/ICSTW52544.2021.00025

Larzelere, R. E., & Huston, T. L. (1980). The dyadic trust scale: toward understanding interpersonal trust in close relationships. *Journal of Marriage and the Family, 42,* 595–604.

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50.30392

Leslie, D. (2019). Understanding Artificial Intelligence Ethics and Safety: A Guide for the Responsible Design and Implementation of AI Systems in the Public Sector. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3403301

Leung, E., Paolacci, G., & Puntoni, S. (2018). Man Versus Machine: Resisting Automation in Identity-Based Consumer Behavior. *Journal of Marketing Research, 55*(6), 818–831. https://doi.org/10.1177/0022243718818423

Leuthesser, L., & Kohli, C. (2015). *Mission Statements and Corporate Identity* (pp. 145–148). https://doi.org/10.1007/978-3-319-13144-3_40

Lewis, J. D., & Weigert, A. J. (1985). Trust as a Social Reality. *Social Forces , 63*, 967–985.

Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., Yi, J., & Zhou, B. (2023). Trustworthy AI: From Principles to Practices. *ACM Computing Surveys*, *55*(9), 1–46.

Li, F., Kashyap, R., Zhou, N., & Yang, Z. (2008). Brand Trust as a Second-order Factor: An Alternative Measurement Model. *International Journal of Market Research*, *50*(6), 817–839. https://doi.org/10.2501/S1470785308200225

Li, Kashyap, R., Zhou, N., & Yang, Z. (2008). Brand Trust as a Second-order Factor: An Alternative Measurement Model. *International Journal of Market Research*, *50*(6), 817–839. https://doi.org/10.2501/S1470785308200225

Li, R., Kumar, A., & Chen, J. H. (2023). How Chatbots and Large Language Model Artificial Intelligence Systems Will Reshape Modern Medicine. *JAMA Internal Medicine*, *183*(6), 596. https://doi.org/10.1001/jamainternmed.2023.1835

Li, W., Yigitcanlar, T., Nili, A., & Browne, W. (2023). Tech Giants' Responsible Innovation and Technology Strategy: An International Policy Review. *Smart Cities*, *6*(6), 3454–3492. https://doi.org/10.3390/smartcities6060153

Li, X., Hess, T. J., & Valacich, J. S. (2008a). Why do we trust new technology? A study of initial trust formation with organizational information systems. *The Journal of Strategic Information Systems*, *17*(1), 39–71. https://doi.org/10.1016/j.jsis.2008.01.001

Li, X., Hess, T. J., & Valacich, J. S. (2008b). Why do we trust new technology? A study of initial trust formation with organizational information systems. *The Journal of Strategic Information Systems*, *17*(1), 39–71. https://doi.org/10.1016/j.jsis.2008.01.001

Liang, Y., & Lee, S. A. (2017). Fear of Autonomous Robots and Artificial Intelligence: Evidence from National Representative Data with Probability Sampling. *International Journal of Social Robotics*, *9*(3), 379–384. https://doi.org/10.1007/s12369-017-0401-3

Liao, Q. V., & Sundar, S. S. (2022). Designing for Responsible Trust in AI Systems: A Communication Perspective. *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1257–1268. https://doi.org/10.1145/3531146.3533182

Liengaard, B. D., Sharma, P. N., Hult, G. T. M., Jensen, M. B., Sarstedt, M., Hair, J. F., & Ringle, C. M. (2021). Prediction: Coveted, Yet Forsaken? Introducing a Cross-Validated

Predictive Ability Test in Partial Least Squares Path Modeling. *Decision Sciences*, *52*(2), 362–392. https://doi.org/10.1111/deci.12445

Lim, K. H., & Benbasat, I. (2000). The Effect of Multimedia on Perceived Equivocality and Perceived Usefulness of Information Systems. *MIS Quarterly*, *24*(3), 449. https://doi.org/10.2307/3250969

Liu, H., Wang, Y., Fan, W., Liu, X., Li, Y., Jain, S., Liu, Y., Jain, A., & Tang, J. (2023). Trustworthy AI: A Computational Perspective. *ACM Transactions on Intelligent Systems and Technology*, *14*(1), 1–59. https://doi.org/10.1145/3546872

Liu, K., & Tao, D. (2022). The roles of trust, personalization, loss of privacy, and anthropomorphism in public acceptance of smart healthcare services. *Computers in Human Behavior*, *127*, 107026. https://doi.org/10.1016/j.chb.2021.107026

Lohmöller, J.-B. (1989). *Latent Variable Path Modeling with Partial Least Squares*. Physica-Verlag HD. https://doi.org/10.1007/978-3-642-52512-4

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, *46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Luhmann, N. (1979). *Trust and Power*. Wiley.

Luhmann, N. (2018). *Trust and Power*. John Wiley & Sons.

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases. *Marketing Science*, mksc.2019.1192. https://doi.org/10.1287/mksc.2019.1192

Lyytinen, K. (1988). Expectation failure concept and systems analysts' view of information system failures: Results of an exploratory study. *Information & Management*, *14*(1), 45–56. https://doi.org/10.1016/0378-7206(88)90066-3

Madhavan, P., & Wiegmann, D. A. (2007). Effects of Information Source, Pedigree, and Reliability on Operator Interaction With Decision Support Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *49*(5), 773–785. https://doi.org/10.1518/001872007X230154

Mahmood, M., Burn, J., Gemoets, L., & Jacquerz, C. (2000). Variables affecting information technology end-user satisfaction: a meta-analysis of the empirical literature.

International Journal of Human-Computer Studies, *52*(4), 751–771. https://doi.org/10.1006/ijhc.1999.0353

Maity, M., & Dass, M. (2014). Consumer decision-making across modern and traditional channels: E-commerce, m-commerce, in-store. *Decision Support Systems*, *61*, 34–46. https://doi.org/10.1016/j.dss.2014.01.008

Maity, M., Dass, M., & Kumar, P. (2018a). The impact of media richness on consumer information search and choice. *Journal of Business Research*, *87*, 36–45. https://doi.org/10.1016/j.jbusres.2018.02.003

Maity, M., Dass, M., & Kumar, P. (2018b). The impact of media richness on consumer information search and choice. *Journal of Business Research*, *87*, 36–45. https://doi.org/10.1016/j.jbusres.2018.02.003

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures*, *90*, 46–60. https://doi.org/10.1016/j.futures.2017.03.006

Malhotra, N. (2007). *Marketing Research: an applied approach* (D. Birks, Ed.; 3rd European Edition). Pearson Education.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995a). An Integrative Model Of Organizational Trust. *Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995b). An Integrative Model Of Organizational Trust. *Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995c). An Integrative Model of Organizational Trust. *The Academy of Management Review*, *20*(3), 709. https://doi.org/10.2307/258792

Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology. *ACM Transactions on Management Information Systems*, *2*(2), 1–25. https://doi.org/10.1145/1985347.1985353

McKnight, D. H., & Chervany, N. (1996). *The Meanings of Trust*.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002a). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, *13*(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002b). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, *13*(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

Medaglia, R., Gil-Garcia, J. R., & Pardo, T. A. (2023). Artificial Intelligence in Government: Taking Stock and Moving Forward. *Social Science Computer Review*, *41*(1), 123–140. https://doi.org/10.1177/08944393211034087

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2022). A Survey on Bias and Fairness in Machine Learning. *ACM Computing Surveys*, *54*(6), 1–35. https://doi.org/10.1145/3457607

Meta. (2021, June 22). *Facebook's five pillars of Responsible AI*. https://ai.meta.com/blog/facebooks-five-pillars-of-responsible-ai/

Microsoft Corporation. (2022). *Microsoft Responsible AI Standard, v2*. https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2022/06/Microsoft-Responsible-AI-Standard-v2-General-Requirements-3.pdf

Miller, K. W., Voas, J., & Laplante, P. (2010). In Trust We Trust. *Computer*, *43*(10), 85–87. https://doi.org/10.1109/MC.2010.289

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, *267*, 1–38. https://doi.org/10.1016/j.artint.2018.07.007

Mirbabaie, M., Brendel, A. B., & Hofeditz, L. (2022). Ethics and AI in Information Systems Research. *Communications of the Association for Information Systems*, *50*(1), 726–753. https://doi.org/10.17705/1CAIS.05034

Mittelstadt, B. (2019). AI Ethics – Too Principled to Fail? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3391293

Modapothula, R., Shaik, N., Arulprakash, P., & Kumar, S. (2022). Study on Potential AI Applications in Childhood Education. *International Journal of Early Childhood Special Education (INT-JECS)*, *14*(3).

Mogavi, R. H., Deng, C., Kim, J. J., Zhou, P., Kwon, Y. D., Metwally, A. H. S., Tlili, A., Bassanelli, S., Bucchiarone, A., Gujar, S., Nacke, L. E., & Hui, P.

(2023). *Exploring User Perspectives on ChatGPT: Applications, Perceptions, and Implications for AI-Integrated Education*.

Mohamad, N. (2020). UNDERSTANDING THE INFLUENCE OF MEDIA RICHNESS IN DEVELOPING CUSTOMER TRUST, COMMITMENT AND LOYALTY. *Journal of Business and Social Development*, *8*(2). https://doi.org/10.46754/jbsd.2020.09.003

Mökander, J., Morley, J., Taddeo, M., & Floridi, L. (2021). Ethics-Based Auditing of Automated Decision-Making Systems: Nature, Scope, and Limitations. *Science and Engineering Ethics*, *27*(4), 44. https://doi.org/10.1007/s11948-021-00319-4

Mondal, S., Das, S., & Vrana, V. G. (2023). How to Bell the Cat? A Theoretical Review of Generative Artificial Intelligence towards Digital Disruption in All Walks of Life. *Technologies*, *11*(2), 44. https://doi.org/10.3390/technologies11020044

Moorman, C., Zaltman, G., & Deshpandé, R. (1992). Relationships between providers and users of market research: the dynamics of trust within and between organizations. *Journal of Marketing Research*, 314–328.

Morgan, R. M., & Hunt, S. D. (1994a). The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing*, *58*(3), 20. https://doi.org/10.2307/1252308

Morgan, R. M., & Hunt, S. D. (1994b). The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing*, *58*(3), 20. https://doi.org/10.2307/1252308

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and Engineering Ethics*, *26*(4), 2141–2168. https://doi.org/10.1007/s11948-019-00165-5

Munn, L. (2023). The uselessness of AI ethics. *AI and Ethics*, *3*(3), 869–877. https://doi.org/10.1007/s43681-022-00209-w

Munuera-Aleman, J. L., Delgado-Ballester, E., & Yague-Guillen, M. J. (2003). Development and Validation of a Brand Trust Scale. *International Journal of Market Research*, *45*(1), 1–18. https://doi.org/10.1177/147078530304500103

Murphy, K., Di Ruggiero, E., Upshur, R., Willison, D. J., Malhotra, N., Cai, J. C., Malhotra, N., Lui, V., & Gibson, J. (2021). Artificial intelligence for good health: a scoping review

of the ethics literature. *BMC Medical Ethics*, *22*(1), 14. https://doi.org/10.1186/s12910-021-00577-8

National Science Foundation. (2024). *Science and Technology: Public Perceptions, Awareness, and Information Sources*. https://ncses.nsf.gov/pubs/nsb20244

Nickel, P. J., Franssen, M., & Kroes, P. (2010). Can We Make Sense of the Notion of Trustworthy Technology? *Knowledge, Technology & Policy*, *23*(3–4), 429–444. https://doi.org/10.1007/s12130-010-9124-6

Niszczota, P., & Kaszás, D. (2020). Robo-investment aversion. *PLOS ONE*, *15*(9), e0239277. https://doi.org/10.1371/journal.pone.0239277

Nunnally, J. C. (1978). *Psychometric theory. Vol. 2*. McGraw-Hill College.

Ofosu-Ampong, K. (2024). Artificial intelligence research: A review on dominant themes, methods, frameworks and future research directions. *Telematics and Informatics Reports*, *14*, 100127. https://doi.org/10.1016/j.teler.2024.100127

Olteanu, A., Castillo, C., Diaz, F., & Kiciman, E. (2016). Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2886526

Paine, K. D. (2003). Guidelines for Measuring Trust in Organizations. In *The Institute for Public Relations*.

Palvia, P., Pinjani, P., Cannoy, S., & Jacks, T. (2011). Contextual constraints in media choice: Beyond information richness. *Decision Support Systems*, *51*(3), 657–670. https://doi.org/10.1016/j.dss.2011.03.006

Peres, R., Schreier, M., Schweidel, D., & Sorescu, A. (2023). On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice. *International Journal of Research in Marketing*, *40*(2), 269–275. https://doi.org/10.1016/j.ijresmar.2023.03.001

Petkovic, D. (2023). It is Not "Accuracy vs. Explainability"—We Need Both for Trustworthy AI Systems. *IEEE Transactions on Technology and Society*, *4*(1), 46–53. https://doi.org/10.1109/TTS.2023.3239921

Pietsch, B. (2021, April 18). *Killed in driverless Tesla car crash, Officials say.* The New York Times. https://www.nytimes.com/2021/04/18/business/tesla-fatal-crash-texas.html

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Prahl, A., & Goh, W. W. P. (2021). "Rogue machines" and crisis communication: When AI fails, how do companies publicly respond? *Public Relations Review*, *47*(4), 102077. https://doi.org/10.1016/j.pubrev.2021.102077

Prem, E. (2023). From ethical AI frameworks to tools: a review of approaches. *AI and Ethics*, *3*(3), 699–716. https://doi.org/10.1007/s43681-023-00258-9

Putnam, R. D. (2000). Bowling alone. *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, 357. https://doi.org/10.1145/358916.361990

Rachels, J., & Rachels, S. (2012). *The Elements of Moral Philosophy* (7th ed.). McGraw-Hill .

Rahwan, I. (2018). Society-in-the-loop: programming the algorithmic social contract. *Ethics and Information Technology*, *20*(1), 5–14. https://doi.org/10.1007/s10676-017-9430-8

Rai, A. (2020). Explainable AI: from black box to glass box. *Journal of the Academy of Marketing Science*, *48*(1), 137–141. https://doi.org/10.1007/s11747-019-00710-5

Reichheld, F. F., & Schefter, P. (2000). E-loyalty: Your secret weapon on the web. . *Harvard Business Review*, 105–113.

Reinhardt, K. (2023). Trust and trustworthiness in AI ethics. *AI and Ethics*, *3*(3), 735–744. https://doi.org/10.1007/s43681-022-00200-5

Riparbelli, V. (2024, January 16). *Educating people about generative AI starts now: Here's where to begin*. World Economic Forum Annual Meeting. https://www.weforum.org/agenda/2024/01/generative-ai-education/

Rodriguez-Rad, C. J., & Ramos-Hidalgo, E. (2018). Spirituality, consumer ethics, and sustainability: the mediating role of moral identity. *Journal of Consumer Marketing*, *35*(1), 51–63. https://doi.org/10.1108/JCM-12-2016-2035

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not So Different After All: A Cross-Discipline View Of Trust. *Academy of Management Review*, *23*(3), 393–404. https://doi.org/10.5465/amr.1998.926617

Ruiz, C. (2019, September 23). *Leading online database to remove 600,000 images after art project reveals its racist bias.* The Art Newspaper. https://www.theartnewspaper.com/2019/09/23/leading-online-database-to-remove-600000-images-after-art-project-reveals-its-racist-bias

Sargeant, A., & Woodliffe, L. (2007). Building Donor Loyalty: The Antecedents and Role of Commitment in the Context of Charity Giving. *Journal of Nonprofit & Public Sector Marketing*, *18*(2), 47–68. https://doi.org/10.1300/J054v18n02_03

Sarstedt, M., Ringle, C. M., Smith, D., Reams, R., & Hair, J. F. (2014). Partial least squares structural equation modeling (PLS-SEM): A useful tool for family business researchers. *Journal of Family Business Strategy*, *5*(1), 105–115. https://doi.org/10.1016/j.jfbs.2014.01.002

Schiffman, S. J., Meile, L. C., & Igbaria, M. (1992). An Examination of End-user Types. *Information and Management*, *22*(4), 207–215.

Schmidt, A., Giannotti, F., Mackay, W., Shneiderman, B., & Väänänen, K. (2021). *Artificial Intelligence for Humankind: A Panel on How to Create Truly Interactive and Human-Centered AI for the Benefit of Individuals and Society* (pp. 335–339). https://doi.org/10.1007/978-3-030-85607-6_32

Schneider, P. A., & Fukuyama, F. (1996). Trust: The Social Virtues and the Creation of Prosperity. *Journal of Marketing*, *60*(3), 129. https://doi.org/10.2307/1251846

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An Integrative Model of Organizational Trust: Past, Present, and Future. *Academy of Management Review*, *32*(2), 344–354. https://doi.org/10.5465/amr.2007.24348410

Schultz, F., & Wehmeier, S. (2010). Institutionalization of corporate social responsibility within corporate communications. *Corporate Communications: An International Journal*, *15*(1), 9–29. https://doi.org/10.1108/13563281011016813

Schwab, K. (2017). *The fourth industrial revolution*. Crown Currency.

Schwartz, S. H. (1992). *Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries* (pp. 1–65). https://doi.org/10.1016/S0065-2601(08)60281-6

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and Abstraction in Sociotechnical Systems. *Proceedings of the Conference on*

Fairness, Accountability, and Transparency, 59–68. https://doi.org/10.1145/3287560.3287598

Sen, S., Raghu, T. S., & Vinze, A. (2009). Demand Heterogeneity in IT Infrastructure Services: Modeling and Evaluation of a Dynamic Approach to Defining Service Levels. *Information Systems Research*, *20*(2), 258–276. https://doi.org/10.1287/isre.1080.0196

Shapiro, S. P. (1987). The Social Control of Impersonal Trust. *American Journal of Sociology*, *93*(3), 623–658. https://doi.org/10.1086/228791

Shareef, M. A., Dwivedi, Y. K., Kumar, V., & Kumar, U. (2017). Content design of advertisement for consumer exposure: Mobile marketing through short messaging service. *International Journal of Information Management*, *37*(4), 257–268. https://doi.org/10.1016/j.ijinfomgt.2017.02.003

Sharma, P. N., Liengaard, B. D., Hair, J. F., Sarstedt, M., & Ringle, C. M. (2023). Predictive model assessment and selection in composite-based modeling using PLS-SEM: extensions and guidelines for using CVPAT. *European Journal of Marketing*, *57*(6), 1662–1677. https://doi.org/10.1108/EJM-08-2020-0636

Sherman, D. K., & Cohen, G. L. (2002). Accepting Threatening Information: Self–Affirmation and the Reduction of Defensive Biases. *Current Directions in Psychological Science*, *11*(4), 119–123. https://doi.org/10.1111/1467-8721.00182

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*, 102551. https://doi.org/10.1016/j.ijhcs.2020.102551

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J.-H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: guidelines for using PLSpredict. *European Journal of Marketing*, *53*(11), 2322–2347. https://doi.org/10.1108/EJM-02-2019-0189

Shneiderman, B. (2020). Bridging the Gap Between Ethics and Practice. *ACM Transactions on Interactive Intelligent Systems*, *10*(4), 1–31. https://doi.org/10.1145/3419764

Shockley-Zalabak, P., & Ellis, K. (2006). The communication of trust. In *The IABC handbook of organizational communication : a guide to internal communication, public relations, marketing and leaderhip.* (pp. 44–55). Jossey-Bass.

Siau, K., & Wang, W. (2018). Building Trust in Artificial Intelligence, Machine Learning, and Robotics. . *Cutter Business Technology Journal*, *31*(2), 47–53.

Siegrist, M., & Cvetkovich, G. (2000). Perception of Hazards: The Role of Social Trust and Knowledge. *Risk Analysis*, *20*(5), 713–720. https://doi.org/10.1111/0272-4332.205064

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., & Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, *550*(7676), 354–359. https://doi.org/10.1038/nature24270

Singh, J., & Sirdeshmukh, D. (2000). Agency and Trust Mechanisms in Consumer Satisfaction and Loyalty Judgments. *Journal of the Academy of Marketing Science*, *28*(1), 150–167. https://doi.org/10.1177/0092070300281014

Sinha, R., & Swearingen, K. (2002). The role of transparency in recommender systems. *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, 830–831. https://doi.org/10.1145/506443.506619

Smith, C. J. (2019). Designing Trustworthy AI: A Human-Machine Teaming Framework to Guide Development. *AAAI FSS-19: Artificial Intelligence in Government and Public Sector*.

Sohn, K., & Kwon, O. (2020). Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products. *Telematics and Informatics*, *47*, 101324. https://doi.org/10.1016/j.tele.2019.101324

Song, J., Baker, J., Lee, S., & Wetherbe, J. C. (2012). Examining online consumers' behavior: A service-oriented view. *International Journal of Information Management*, *32*(3), 221–231. https://doi.org/10.1016/j.ijinfomgt.2011.11.002

Sousa, W. G. de, Melo, E. R. P. de, Bermejo, P. H. D. S., Farias, R. A. S., & Gomes, A. O. (2019). How and where is artificial intelligence in the public sector going? A literature review and research agenda. *Government Information Quarterly*, *36*(4), 101392. https://doi.org/10.1016/j.giq.2019.07.004

Stamboliev, E., & Christiaens, T. (2024). How "empty" is Trustworthy AI? A discourse analysis of the Ethics Guidelines of Trustworthy AI. *Critical Policy Studies*, 1–18. https://doi.org/10.1080/19460171.2024.2315431

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society*, *36*(2), 111–147.

Strobel, G., Banh, L., Möller, F., & Schoormann, T. (2024). Exploring Generative Artificial Intelligence: A Taxonomy and Types. *Proceedings of the 57th Hawaii International Conference on System Sciences*, Page 4546-45555.

Suh, K. S. (1999). Impact of communication medium on task performance and satisfaction: an examination of media-richness theory. *Information & Management*, *35*(5), 295–312. https://doi.org/10.1016/S0378-7206(98)00097-4

Sujan, M., Furniss, D., Hawkins, R., & Habli, I. (2020). *Human Factors of Using Artificial Intelligence in Healthcare: Challenges That Stretch Across Industries.* .

Szajna, B., & Scamell, R. W. (1993). The Effects of Information System User Expectations on Their Performance and Perceptions. *MIS Quarterly*, *17*(4), 493. https://doi.org/10.2307/249589

Taddeo, M., & Floridi, L. (2018a). How AI can be a force for good. *Science*, *361*(6404), 751–752. https://doi.org/10.1126/science.aat5991

Taddeo, M., & Floridi, L. (2018b). How AI can be a force for good. *Science*, *361*(6404), 751–752. https://doi.org/10.1126/science.aat5991

Taylor, J. E. T., & Taylor, G. W. (2021). Artificial cognition: How experimental psychology can help generate explainable artificial intelligence. *Psychonomic Bulletin & Review*, *28*(2), 454–475. https://doi.org/10.3758/s13423-020-01825-5

The Economist. (2017). *The world's most valuable resource is no longer oil, but data*. https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data

The White House. (2023). *NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN 2023 UPDATE*. https://www.whitehouse.gov/wp-content/uploads/2023/05/National-Artificial-Intelligence-Research-and-Development-Strategic-Plan-2023-Update.pdf#:~:text=URL%3A%20https%3A%2F%2Fwww.whitehouse.gov%2Fwp

Thiebes, S., Lins, S., & Sunyaev, A. (2021). Trustworthy artificial intelligence. *Electronic Markets*, *31*(2), 447–464. https://doi.org/10.1007/s12525-020-00441-4

Thomas, T., Singh, L., & Gaffar, K. (2013). The utility of the UTAUT model in explaining mobile learning adoption in higher education in Guyana. *International Journal of Education and Development Using ICT*, *9*(3). https://www.learntechlib.org/p/130274/

Toreini, E., Aitken, M., Coopamootoo, K., Elliott, K., Zelaya, C. G., & van Moorsel, A. (2020). The relationship between trust in AI and trustworthy machine learning technologies. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 272–283. https://doi.org/10.1145/3351095.3372834

Trawnih, A., Al-Masaeed, S., Alsoud, M., & Alkufahy, A. M. (2022). Understanding artificial intelligence experience: A customer perspective. *International Journal of Data and Network Science*, *6*(4), 1471–1484. https://doi.org/10.5267/j.ijdns.2022.5.004

Treviño, L. K., Webster, J., & Stein, E. W. (2000). Making Connections: Complementary Influences on Communication Media Choices, Attitudes, and Use. *Organization Science*, *11*(2), 163–182. https://doi.org/10.1287/orsc.11.2.163.12510

Trevino, L., Lengel, R., & Daft, R. (1987). Media Symbolism, Media Richness, and Media Choice in Organizations. *Communication Research*, *14*(5), 553–574. https://doi.org/10.1177/009365087014005006

Tseng, C.-H., & Wei, L.-F. (2020). The efficiency of mobile media richness across different stages of online consumer behavior. *International Journal of Information Management*, *50*, 353–364. https://doi.org/10.1016/j.ijinfomgt.2019.08.010

Turner, M., Kitchenham, B., Brereton, P., Charters, S., & Budgen, D. (2010). Does the technology acceptance model predict actual use? A systematic literature review. *Information and Software Technology*, *52*(5), 463–479. https://doi.org/10.1016/j.infsof.2009.11.005

Vakkuri, V., Kemell, K.-K., Kultanen, J., & Abrahamsson, P. (2020). The Current State of Industrial Practice in Artificial Intelligence Ethics. *IEEE Software*, *37*(4), 50–57. https://doi.org/10.1109/MS.2020.2985621

van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics*, *1*(3), 213–218. https://doi.org/10.1007/s43681-021-00043-6

Varshney, K. R. (2019). Trustworthy machine learning and artificial intelligence. *XRDS: Crossroads, The ACM Magazine for Students*, *25*(3), 26–29. https://doi.org/10.1145/3313109

Venkatesh, Morris, Davis, & Davis. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425. https://doi.org/10.2307/30036540

Vereschak, O., Bailly, G., & Caramiaux, B. (2021). How to Evaluate Trust in AI-Assisted Decision Making? A Survey of Empirical Methodologies. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–39. https://doi.org/10.1145/3476068

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science*, *44*(1), 119–134. https://doi.org/10.1007/s11747-015-0455-4

Wacksman, J. (2021). Digitalization of contact tracing: balancing data privacy with public health benefit. *Ethics and Information Technology*, *23*(4), 855–861. https://doi.org/10.1007/s10676-021-09601-2

Wafa, A. W. A., & Muzammil Hussain, M. H. (2021). A Literature Review of Artificial Intelligence. *UMT Artificial Intelligence Review*, *1*(1), 1–1. https://doi.org/10.32350/AIR.11.01

Wang, C., Teo, T. S. H., & Janssen, M. (2021). Public and private value creation using artificial intelligence: An empirical study of AI voice robot users in Chinese public sector. *International Journal of Information Management*, *61*, 102401. https://doi.org/10.1016/j.ijinfomgt.2021.102401

Wang, H., Huang, J., & Zhang, Z. (2019). *The Impact of Deep Learning on Organizational Agility*.

Wang, Y. D., & Emurian, H. H. (2005). An overview of online trust: Concepts, elements, and implications. *Computers in Human Behavior*, *21*(1), 105–125. https://doi.org/10.1016/j.chb.2003.11.008

Wang, Z. (2022). Media Richness and Continuance Intention to Online Learning Platforms: The Mediating Role of Social Presence and the Moderating Role of Need for Cognition. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.950501

Wood, S. (2024). The Future of Consumer Research Methods: Lessons of a Prospective Retrospective. *Journal of Consumer Research*, *51*(1), 151–156. https://doi.org/10.1093/jcr/ucae017

Wright, J. (2024). The Development of AI Ethics in Japan: Ethics-washing Society 5.0? *East Asian Science, Technology and Society: An International Journal*, *18*(2), 117–134. https://doi.org/10.1080/18752160.2023.2275987

Xia, W., Zhou, D., Xia, Q., & Zhang, L. (2020). Design and implementation path of intelligent transportation information system based on artificial intelligence technology. *The Journal of Engineering*, *2020*(13), 482–485. https://doi.org/10.1049/joe.2019.1196

Xie, X., Wu, Y., & Blanco-Gonzalez Tejerob, C. (2024). How Responsible Innovation Builds Business Network Resilience to Achieve Sustainable Performance During Global Outbreaks: An Extended Resource-Based View. *IEEE Transactions on Engineering Management*, 1–15. https://doi.org/10.1109/TEM.2022.3186000

Yi, M. Y., Jackson, J. D., Park, J. S., & Probst, J. C. (2006). Understanding information technology acceptance by individual professionals: Toward an integrative view. *Information & Management*, *43*(3), 350–363. https://doi.org/10.1016/j.im.2005.08.006

Yingprayoon, J. (2023). *Public Understanding of Science and Technology* (pp. 181–190). https://doi.org/10.1007/978-3-031-24259-5_13

Zerilli, J., Bhatt, U., & Weller, A. (2022). How transparency modulates trust in artificial intelligence. *Patterns*, *3*(4), 100455. https://doi.org/10.1016/j.patter.2022.100455

Zhang, B., & Dafoe, A. (2019). Artificial Intelligence: American Attitudes and Trends. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3312874

Zhang, Y., & Gosline, R. (2023a). Human favoritism, not AI aversion: People's perceptions (and bias) toward generative AI, human experts, and human–GAI collaboration in persuasive content generation. *Judgment and Decision Making*, *18*, e41. https://doi.org/10.1017/jdm.2023.37

Zhang, Y., & Gosline, R. (2023b). Human favoritism, not AI aversion: People's perceptions (and bias) toward generative AI, human experts, and human–GAI collaboration in persuasive content generation. *Judgment and Decision Making*, *18*, e41. https://doi.org/10.1017/jdm.2023.37

Zhanga, Z. T., & Hußmann, H. (2021, July). How to Manage Output Uncertainty: Targeting the Actual End User Problem in Interactions with AI. *Joint Proceedings of the ACM IUI 2021 Workshops - Co- Located with the 26th ACM Conference on Intelligent User Interfaces*.

Zhou, J., & Chen, F. (2023). AI ethics: from principles to practice. *AI & SOCIETY*, *38*(6), 2693–2703. https://doi.org/10.1007/s00146-022-01602-z

Ziegler, J., & Donkers, T. (2024). From explanations to human-AI co-evolution: charting trajectories towards future user-centric AI. *I-Com*, *23*(2), 263–272. https://doi.org/10.1515/icom-2024-0020

Zucker, L. G. (1986). Production of trust: Institutional sources of economic structure. *Research in Organizational Behavior*, *8*, 53–111.

Annex

Annex A – Questionnaire

Dear Participant,

My name is Alina, a student at the ISCTE Business School of Lisbon, currently working on my master's thesis. I want to explore how we can align technological progress and human values.

For that, I need your help to understand what holds significance to you when it comes to generative AI.

Among all participants, three winners will be drawn to receive 20€ vouchers for Amazon, Spotify, or AirBnB (free choice). At the end of the study, you will find a link to participate in the prize draw.

Please remember the following points:

There are no right or wrong answers; your personal opinions and experiences are what matter.

All data collected will be stored and analyzed anonymously.

Your responses cannot be traced back to you and will be handled confidentially.

I encourage you to answer each question spontaneously and honestly.

I care about the quality of my survey data. For me to get the most accurate measures of your opinions, it is important that you provide thoughtful answers to each question in this survey. The survey should only take about 10 – 15 minutes, so I kindly ask you to read each question thoroughly.

Question 1: Do you commit to providing thoughtful answers to the questions in this survey?

Yes, I will	
No, I will not (do not want to participate in the study	

Should you have any questions or comments, please contact me via email at afral@iscte-iul.pt.

Thank you for your contribution to my work!

Best,

Alina

BEFORE WE START

To adhere to the standards of empirical research, I require your consent.

Voluntary

Your participation in this study is entirely optional. You have the freedom to withdraw from the study at any point without facing any negative consequences.

Anonymity

Your information will be handled confidentially, analyzed only in anonymous formats, and will not be shared with third parties. Demographic details such as age or gender will not be utilized for identification purposes.

Question 2: I agree to the processing of my personal data in accordance with the information provided herein.

Yes	
No (do not want to participate in the study	

In the following, you will be shown a company statement from Generative AI (GAI) company *Elle Media*.

Background Information:

GAI *is a type of artificial intelligence (AI) that specializes in creating new content or data rather than just processing existing information. Unlike regular AI, which follows set rules, GAI can produce original outputs based on what it has learned. GAI can create realistic images, generate natural language text, compose music, and even produce entirely new designs or ideas autonomously. An example of a GAI technology is OpenAI's ChatGPT, which can generate human-like text, enabling tasks like language translation and content creation.*

Your Task: Please read the whole document carefully. Based on the document, you will be asked about your perceptions.

Please note: There are no right or wrong answers. This study is only about your individual assessment.

[show stimulus material: according to testgroup - either (1) low, (2) moderate, (3) high MR]

How do you rate the Media Richness of this document?

Media Richness refers to a communication medium's effectiveness in providing information that meets users' needs to ensure alignment with audience expectations and requirements.

Question 3: The document provides a wide range of information.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 4: The document provides in-depth information.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 5: Please select „agree“ on the scale to indicate that you are paying attention. [*attention check*]

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 6: The document meets my requirements for information content.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 7: The document provides high-quality information.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 8: The document provides trustworthy information.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 9: The document provided by the company can help me.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Next, I would like to hear your impression of the GAI technology offered by the company.
How much do you agree or disagree with the following statements?

Question 10: The offered GAI technology cares about our well-being.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 11: The offered GAI technology is sincerely concerned about addressing the problems of human users.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 12: The offered GAI technology tries to be helpful and do not operate out of selfish interest.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 13: The offered GAI technology is truthful in their dealings.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 14: The offered GAI technology keeps their commitments and delivers on its promises.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 15: The offered GAI technology is honest and do not abuse the information and advantage they have over its users.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 16: The offered GAI technology has the features necessary to complete key tasks.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 17: The offered GAI technology is competent in its area of expertise.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 18 The offered GAI technology is reliable.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 19: The offered GAI technology is dependable.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Now I'd like to hear your opinion regarding the company.

To what extent do you agree or disagree with the following statements?

Question 20: This company does a good job.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 21: I expect the company to deliver on its promise.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 22: I am confident in the company's ability to perform well.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 23: The quality of this company seems to be very consistent.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 24: The company has good intentions towards its costumers.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 25: It will respond constructively if I have any product-related problems.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 26: It would do its best to help me if I had a problem.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 27: It cares about my needs.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Question 28: This company gives me a sense of security.

1 - Strongly disagree	2	3	4	5 - Strongly agree

Last questions!

How important are these values in the design of GAI technologies that interact with us?

Question 29: Accountability - AI at any step is accountable for considering the system's impact in the world.

1 – Not at all important	2	3	4	5 – Very Important

Question 30: Transparency – Transparency requirements that reduce the opacity of systems.

1 – Not at all important	2	3	4	5 – Very Important

Question 31: Privacy and data governance – Competent authorities who implement legal frameworks and guidelines for testing and certification of GAI-enabled products and services.

1 – Not at all important	2	3	4	5 – Very Important

Question **32**: Diversity, non-discrimination and fairness: The application of rules designed to protect fundamental human rights, such as equality.

1 – Not at all important	2	3	4	5 – Very Important

Question **33**: Technical robustness and safety: Systems are developed in a responsible manner with proper consideration of risks.

1 – Not at all important	2	3	4	5 – Very Important

Question **34**: Human agency and oversight: Human oversight and control throughout the lifecycle of GAI products.

1 – Not at all important	2	3	4	5 – Very Important

Question **35**: Societal and environmental well-being: AI systems that conform to the best standards of sustainability and address like issues climate change and environmental justice.

1 – Not at all important	2	3	4	5 – Very Important

To finish, a few questions about yourself:

There's no need to be concerned about the data being linked to you. If you prefer to withhold certain information, select the option "do not specify".

Question 36: How old are you?

Please enter your age in years.

	years

Question 37: Which gender do you feel you belong to?

Male	
Female	
Divers	
Do not specify	

Question 38: What level of education do you have?

Please select the highest qualification you have achieved so far.

No schooling completed	
Basic Education (9th grade)	
High School (12th grade)	
Bachelor's Degree	
Post-Graduation	
Master's Degree	
PhD	
Other qualification, namely	
Do not specify	

Question 39: What describes best your current situation?

Student	
Working student	
Worker	
Unemployed	
Retired	
Other situation	
Do not specify	

Question 40: Approximately, what is your monthly income?

This refers to the sum you have at your disposal each month (after taxes and social insurance deductions) for your living expenses, regardless of whether the funds come from employment, relatives, or other sources.

I have no income	
0-249 €	
250 – 499 €	
500 - 999 €	
1000 – 1499 €	
1500 – 1999 €	
2000 – 2499 €	
2500 – 2999 €	
3000 € or more	

Your answers have been saved ✓

Thank you so much for your participation!

If you wish to participate in the prize draw, you will be redirected to another site to ensure that your survey answers remain anonymous.

Please click here to proceed.

Annex B – Measurement of items

Table X *- Items for measurement of the constructs*

	Item	Scale	References
Values			
Value 1	How important are these values in the design of GAI technologies that interact with us? [Privacy and data governance: Competent authorities who implement legal frameworks and guidelines for testing and certification of AI-enabled products and services.]	1 (not at all important) – 5 (very important)	Choung et al. (2023a); David et al. (2024)
Value 2	How important are these values in the design of GAI technologies that interact with us? [Human agency and oversight: Human oversight and control throughout the lifecycle of AI products.]		
Value 3	How important are these values in the design of GAI technologies that interact with us? [Technical robustness and safety: Systems are developed in a responsible manner with proper consideration of risks.]		
Value 4	How important are these values in the design of GAI technologies that interact with us? [Transparency: Transparency requirements that reduce the opacity of systems.]		
Value 5	How important are these values in the design of GAI technologies that interact with us? [Diversity, non-discrimination, and fairness: The application of rules designed to protect fundamental human rights, such as equality.]		
Value 6	How important are these values in the design of GAI technologies that interact with us? [Societal and environmental well-being: AI systems that conform to the best standards of sustainability and address like issues climate change and environmental justice.]		
Value 7	How important are these values in the design of GAI technologies that interact with us? [Accountability: AI at any step is accountable for considering the system's impact in the world.]		

Table X *- Items for measurement of the constructs*

	Item	Scale	References
Generative AI Trust			
GT 1	The offered GAI technology cares about our well-being. *(Benevolence 1)*	1 (strongly disagree) – 5 (strongly agree)	David et al. (2024)
GT 2	The offered GAI technology is sincerely concerned about addressing the problems of human users. *(Benevolence 2)*		
GT 3	The offered GAI technology tries to be helpful and do not operate out of selfish interest. *(Benevolence 3)*		
GT 4	The offered GAI technology is truthful in their dealings. *(Integrity 1)*		
GT 5	The offered GAI technology keeps their commitments and delivers on its promises. *(Integrity 2)*		
GT 6	The offered GAI technology is honest and do not abuse the information and advantage they have over its users. *(Integrity 3)*		
GT 7	The offered GAI technology works well. (*Competence 1)*		
GT 8	The offered GAI technology has the features necessary to complete key tasks. *(Competence 2)*		
GT 9	The offered GAI technology is competent in its area of expertise. *(Competence 3)*		
GT 10	The offered GAI technology is reliable. *(Competence 4)*		
GT 11	The offered GAI technology is dependable. *(Competence 5)*		
Brand Trust			
BT 1	This brand does a good job. *(Competence 1)*	1 (strongly disagree) – 5 (strongly agree)	Li et al. (2008)
BT 2	I expect the brand to deliver on its promise. *(Competence 2)*		
BT 3	I am confident in the brand's ability to perform well. (*Competence 3)*		

Table X *- Items for measurement of the constructs*

	Item	Scale	References
BT 4	The quality of this brand seems to be very consistent. *(Competence 4)*	1 (strongly disagree) – 5 (strongly agree)	Li et al. (2008)
BT 5	The brand has good intentions towards its customers. *(Benevolence 1)*		
BT 6	It will respond constructively if I have any product-related problems. *(Benevolence 2)*		
BT 7	It would do its best to help me if I had a problem. *(Benevolence 3)*		
BT 8	It cares about my needs. *(Benevolence 4)*		
BT 9	This brand gives me a sense of security. *(Benevolence 5)*		
Media Richness			
MR 1	The mission statement provides a wide range of information. *(Content Richness 1)*	1 (strongly disagree) – 5 (strongly agree)	Z. Wang (2022)
MR 2	The mission statement provides in-depth information. *(Content Richness 2)*		
MR 3	The mission statement meets my requirements for information content. *(Content Richness 3)*		
MR 4	The mission statement provides high-quality information. *(Quality Richness 1)*		
MR 5	The mission statement provides trustworthy information. *(Quality Richness 2)*		
MR 6	The mission statement provided by the brand can help me. *(Quality Richness 3)*		

Annex C– Preliminary Test Results

Table B4 *- Sociodemographic characteristics of the Preliminary Study*

	n	**%**	**M**	**SD**
Age			31.2	11
Gender				
Male	8	38.1		
Female	13	61.9		
Diverse	0	0		
Education				
No schooling completed	0	0		
Basic education. (9th grade)	2	9.5		
High school (12th grade)	3	14.3		
Bachelor's degree	10	47.6		
Post Graduation	1	4.8		
Master's Degree	4	19.0		
PhD	1	4.8		
Income				
No income	1	4.8		
0 – 249 €	1	3.4		
250 – 499 €	1	10.3		
500 - 999 €	7	24.1		
1000 – 1499 €	3	13.8		
1500 – 1999 €	4	19.0		
2000 – 2499 €	2	9.5		
2500 – 2999 €	1	4.8		
3000 € or more	1	4.8		
Current Situation				
Student	3	14.3		
Working Student	8	38.1		
Worker	10	47.6		
Unemployed	0	0		
Other situation	0	0		

Notes. *N*= 21

Table X *- Construct reliability and validity of the Preliminary Study*

Cons	**Cronbach's alpha**	**Composite reliability (rho_a)**	**Composite reliability (rho_c)**	**Average variance extracted (AVE)**
MR	.876	.901	.908	.628
BT	.843	.869	.877	.461
GT	.936	.951	.944	.629

Notes. $N = 21$

Annex D – Sample Characterization Results

Table x *- Education frequency*

Education	Frequency	Percent	Valid Percent	Cumulative Percent
Basic Education (9th grade)	4	1.3	1.3	1.3
High School (12th grade)	45	15	15	16.3
Bachelor Degree	154	51.3	51.3	67.7
Post-Graduation	25	8.3	8.3	76
Masters Degree	55	18.3	18.3	94.3
PhD	6	2	2	96.3
Other qualification	7	2,3	2,3	98.7
Do not specify	4	1.3	1.3	100
Total	300	100	100	

***Table x** - Income frequency*

Income	Frequency	Percent	Valid Percent	Cumulative Percent
I have no income	24	8	8	8
0 - 249 €	61	20.3	20.3	28.3
250 - 499 €	31	10.3	10.3	38.7
500 - 999€	43	14.3	14.3	53
1000 - 1499 €	39	13	13	66
1500 - 1999 €	23	7.7	7.7	73.7
2000 - 2499 €	34	11.3	11.3	85
2500 - 2999€	22	7.3	7.3	92.3
3000 € or more	23	7.7	7.7	100
Total	300	100	100	

***Table x** - Situation frequency*

Situation	Frequency	Percent	Valid Percent	Cumulative Percent
Student	21	7	7	7
Working Student	28	9.3	9.3	16.3
Worker	184	61,3	61.3	77.7
Unemployed	38	12.7	12,7	90.3
Retired	3	1	1	91.3
Other situation	24	8	8	99.3
Do not specify	2	0.7	0.7	100
Total	300	100	100	

Annex E – Preliminary Control Checks

Tabelle X - Construct Reliability & Validity Pre-Test

Construct	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
GT	0.936	0.968	0.942	0.622
BT	0.843	0.869	0.875	0.458
MR	0.876	0.899	0.908	0.627

Notes. $N = 21$

***Tabelle X** - ANOVA Oneway Outputs: MR Descriptives*

					95% Confidence Interval for Mean			Max
Testgroup	**N**	**Mean**	**SD**	**Std. Error**	**Lower Bound**	**Upper Bound**	**Min**	
Low MR	102	3.4183	.81337	.08054	3.2585	3.5781	1.17	5.00
Moderate	102	3.8252	.67827	.06716	3.6919	3.9584	1.83	5.00
High	96	4.0677	.55492	.05664	3.9553	4.1801	2.83	5.00
Total	300	3.7644	.74053	.04275	3.6803	3.8486	1.17	5.00

Tabelle X - ANOVA Oneway Outputs: Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
Media Richness	Based on Mean	9.383	2	297	<.001
	Based on Median	8.942	2	297	<.001
	Based on Median and with adjusted df	8.942	2	285,368	<.001
	Based on trimmed mean	9.249	2	297	<.001

Tabelle X - ANOVA Oneway Outputs: MR

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	21.426	2	10.713	22.322	<.001
Within Groups	142.539	297	.480		
Total	163.965	299			

***Tabelle X** - ANOVA Oneway Outputs: Effect Sizes[a]*

			95% Confidence Interval	
		Point Estimate	Lower	Upper
Media Richness	Eta-squared	.131	.065	.199
	Epsilon-squared	.125	.058	.194
	Omega-squared Fixed-effect	.124	.058	.193
	Omega-squared Random-effect	.066	.030	.107

Notes. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model

Table X: Tukey's Honestly Significant Difference (HSD) Post Hoc Test

Testgroup	***N***	**1**	**2**	**3**
Low MR	102	3.4183		
Moderate MR	102		3.8252	
High MR	96			4.0677

Notes. Means for groups in homogeneous subsets are displayed. a. Uses Harmonic Mean Sample Size = 99,918. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed. *Subset for alpha = 0.05*

Table X *- Principal Component Analysis: Total Variance Explained*

Component	Initial Eigenvalues		Extraction Sums of Squared Loadings			
Total	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**	
1	13.295	41.548	41.548	13.295	41.548	41.548
2	3.138	9.807	51.354	3.138	9.807	51.354
3	1.588	4.963	56.318	1.588	4.963	56.318
4	1.018	3.181	59.498	1.018	3.181	59.498
5	0.938	2.93	62.428			
6	0.771	2.408	64.837			
7	0.762	2.38	67.217			
8	0.744	2.324	69.541			
9	0.691	2.159	71.7			
10	0.666	2.081	73.781			
11	0.64	1.999	75.78			
12	0.609	1.903	77.684			
13	0.558	1.745	79.428			
14	0.546	1.705	81.133			
15	0.494	1.545	82.678			
16	0.466	1.457	84.135			
17	0.44	1.375	85.51			
18	0.433	1.353	86.863			
19	0.428	1.336	88.199			
20	0.411	1.286	89.484			
21	0.38	1.187	90.671			
22	0.368	1.149	91.821			
23	0.343	1.071	92.891			
24	0.33	1.032	93.923			
25	0.303	0.947	94.869			
26	0.284	0.888	95.758			
27	0.262	0.818	96.576			
28	0.241	0.752	97.328			
29	0.237	0.739	98.068			

30	0.228	0.713	98.78
31	0.213	0.667	99.447
32	0.177	0.553	100

Annex F - PLS Algorithm Results

Table xx – Model Fit

	Saturated model	**Estimated model**
SRMR	0.061	0.099
d_ULS	1.731	4.567
d_G	0.558	0.704
Chi-square	908.695	1062.741
NFI	0.839	0.811

Table x *- Multicollinearity Statistics (VIF)*

Item	VIF
BT 1	2.190
BT 2	1.750
BT 3	2.225
BT 4	2.271
BT 5	2.316
BT 6	2.154
BT 7	2.797
BT 8	2.344
BT 9	2.151
GT 1	2.195
GT 2	2.654
GT 3	2.017
GT 4	2.259
GT 5	1.979
GT 6	2.462
GT 8	1.686
GT 9	2.499
GT 10	2.031
MR 1	1.821
MR 2	1.882
MR 3	1.919
MR 4	2.893
MR 5	1.975
MR 6	1.706
Values 1	1.560
Values 2	1.667
Values 3	1.757

Values 4	1.451
Values 5	1.586
Values 6	1.422

Annex G – Bootstrapping Results

Table x – Outer Loadings and p values

	Original sample	Sample mean	SD	T statistics	P values
BT_1 <- BT	0.797	0.796	0.026	31.018	0.000
BT_2 <- BT	0.711	0.708	0.040	17.993	0.000
BT_3 <- BT	0.798	0.798	0.024	33.794	0.000
BT_4 <- BT	0.801	0.800	0.024	33.851	0.000
BT_5 <- BT	0.797	0.796	0.027	30.001	0.000
BT_6 <- BT	0.768	0.767	0.030	25.224	0.000
BT_7 <- BT	0.831	0.830	0.022	38.459	0.000
BT_8 <- BT	0.798	0.797	0.025	32.099	0.000
BT_9 <- BT	0.758	0.758	0.031	24.456	0.000
GT 1 <- GT	0.759	0.757	0.030	25.566	0.000
GT 10 <- GT	0.718	0.717	0.042	17.079	0.000
GT 2 <- GT	0.810	0.809	0.022	37.031	0.000
GT 3 <- GT	0.762	0.761	0.034	22.278	0.000
GT 4 <- GT	0.782	0.780	0.029	26.716	0.000
GT 5 <- GT	0.747	0.746	0.038	19.639	0.000
GT 6 <- GT	0.816	0.814	0.025	32.052	0.000
GT 8 <- GT	0.721	0.721	0.031	23.285	0.000
GT 9 <- GT	0.798	0.797	0.030	26.635	0.000
MR 1 <- MR	0.735	0.734	0.044	16.845	0.000

MR 2 <- MR	0.750	0.749	0.032	23.517	0.000
MR 3 <- MR	0.787	0.786	0.027	29.486	0.000
MR 4 <- MR	0.876	0.875	0.015	58.728	0.000
MR 5 <- MR	0.797	0.797	0.020	40.299	0.000
MR 6 <- MR	0.755	0.754	0.031	24.266	0.000
Value 1 <- Values	0.728	0.725	0.041	17.781	0.000
Values 2 <- Values	0.667	0.658	0.058	11.460	0.000
Values 3 <- Values	0.732	0.725	0.048	15.112	0.000
Values 4 <- Values	0.760	0.763	0.035	21.945	0.000
Values 5 <- Values	0.723	0.720	0.039	18.342	0.000
Values 6 <- Values	0.686	0.682	0.047	14.740	0.000

Table x – Path coefficients and p values

	Original sample	**Sample mean**	**SD**	**T statistics**	**P values**
MR -> BT	0.712	0.712	0.034	21.014	0.000
MR -> GT	0.685	0.686	0.033	20.594	0.000
MR -> Values	0.291	0.299	0.054	5.360	0.000
Values -> BT	0.100	0.103	0.047	2.152	0.031
Values -> GT	0.113	0.114	0.049	2.310	0.021

Table 33 – Specific indirect effects (complete)

	Original sample	Sample mean	SD	T statistics	P values
MR -> Values -> BT	0.029	0.031	0.016	1.849	0.065
MR -> Values -> GT	0.033	0.034	0.016	2.064	0.039

Table XX – Total indirect effects

	Original sample	Sample mean	SD	T statistics	P values
MR -> BT	0.029	0.031	0.016	1.849	0.065
MR -> GT	0.033	0.034	0.016	2.064	0.039

Table 37 – Total effects

	Original sample	Sample mean	SD	T statistics	P values
MR -> BT	0.741	0.743	0.028	26.808	0.000
MR -> GT	0.717	0.720	0.028	26.069	0.000
MR -> Values	0.291	0.299	0.054	5.360	0.000
Values -> BT	0.100	0.103	0.047	2.152	0.031
Values -> GT	0.113	0.114	0.049	2.310	0.021

Table 37 – Q²Predict values

	Q²predict
BT_1_x	0.422
BT_2_x	0.225
BT_3_x	0.394
BT_4_x	0.342
BT_5_x	0.317
BT_6_x	0.278
BT_7_x	0.331
BT_8_x	0.361
BT_9_x	0.305
GT_10_x	0.261
GT_1_x	0.291
GT_2_x	0.347
GT_3_x	0.242
GT_4_x	0.256
GT_5_x	0.282
GT_6_x	0.300
GT_8_x	0.373
GT_9_x	0.301
v_28_x	0.041
v_29_x	0.006
v_30_x	0.007
v_31_x	0.094
v_32_x	0.028
v_35_x	0.023

Table 37 – PLS-SEM vs. Indicator average (IA)

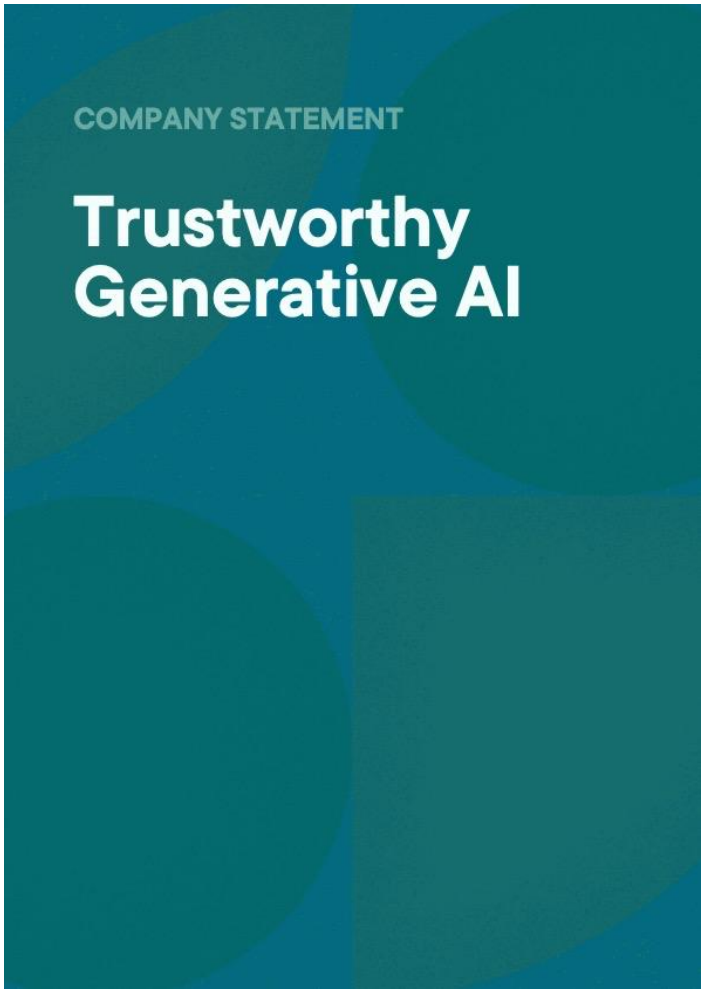
	PLS loss	**IA loss**	**Average loss difference**	**t value**	**p value**
BT	0.475	0.709	-0.234	6.653	0.000
GT	0.507	0.719	-0.212	6.606	0.000
Values	0.860	0.891	-0.031	1.924	0.055
Overall	0.583	0.758	-0.175	7.008	0.000

Table 37 – PLS-SEM vs. Linear Model (LM)

	PLS loss	**LM loss**	**Average loss difference**	**t value**	**p value**
BT	0.475	0.467	0.007	0.721	0.471
GT	0.507	0.499	0.007	0.726	0.468
Values	0.860	0.870	-0.010	1.023	0.307
Overall	0.583	0.580	0.003	0.403	0.688

Annex H - Stimuli

Figuret 9 *– Stimuli Testgroup 1, low MR*



Our Quality Promise

1. **Privacy & Data Governance**
   We prioritize privacy and data protection
2. **Human Agency & Oversight**
   We ensure human oversight in our GAI systems
3. **Technical Robustness & Safety**
   We prioritize safety and reliability
4. **Transparency**
   Our GAI systems are transparent
5. **Diversity, Non-discrimination, Fairness**
   We promote diversity and fairness
6. **Societal & Environmental Well-Being**
   We consider societal and environmental impact
7. **Accountability**
   We uphold accountability in our GAI systems

Figure 10 *– Stimuli Testgroup 2, moderate MR*

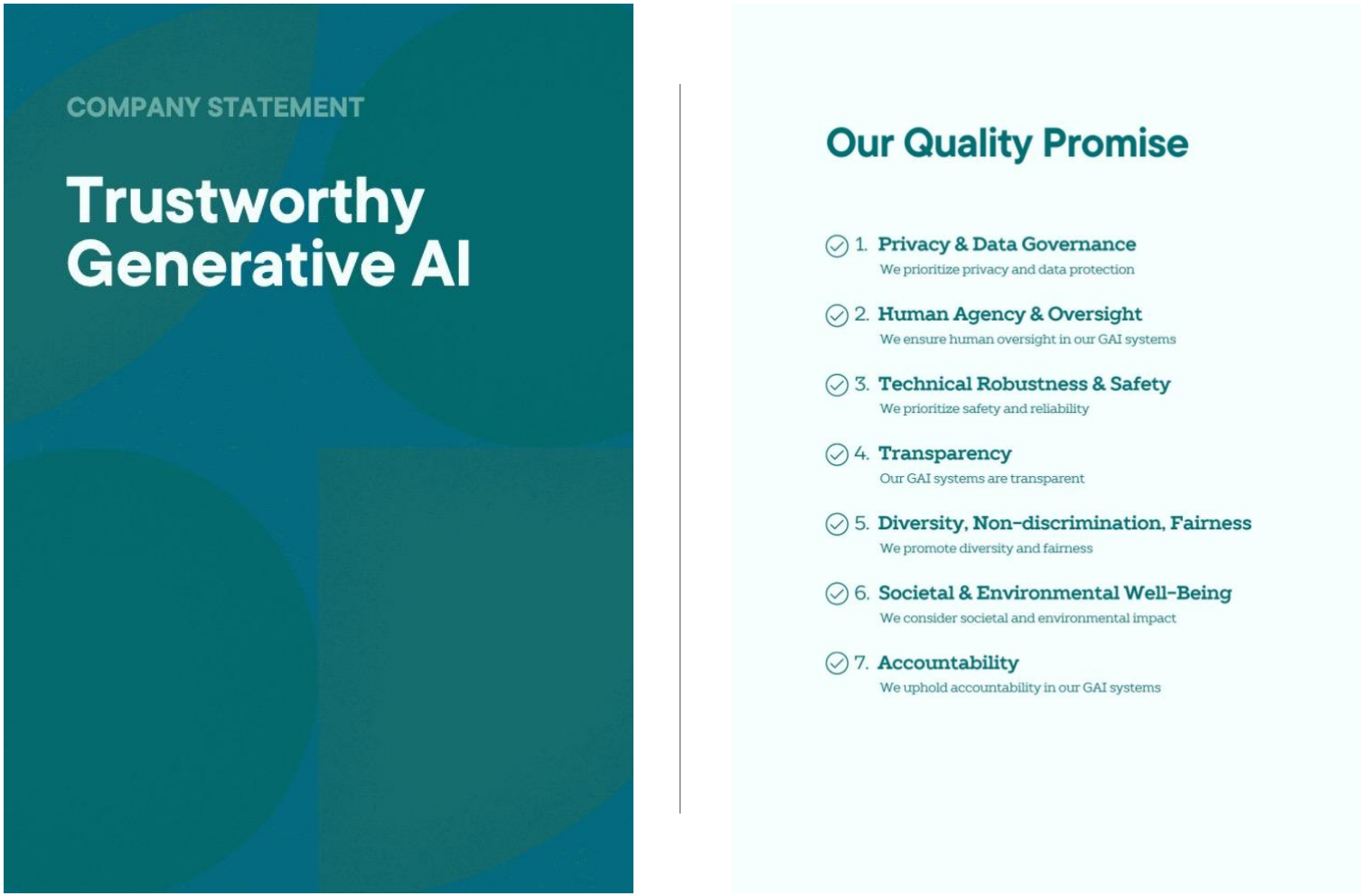
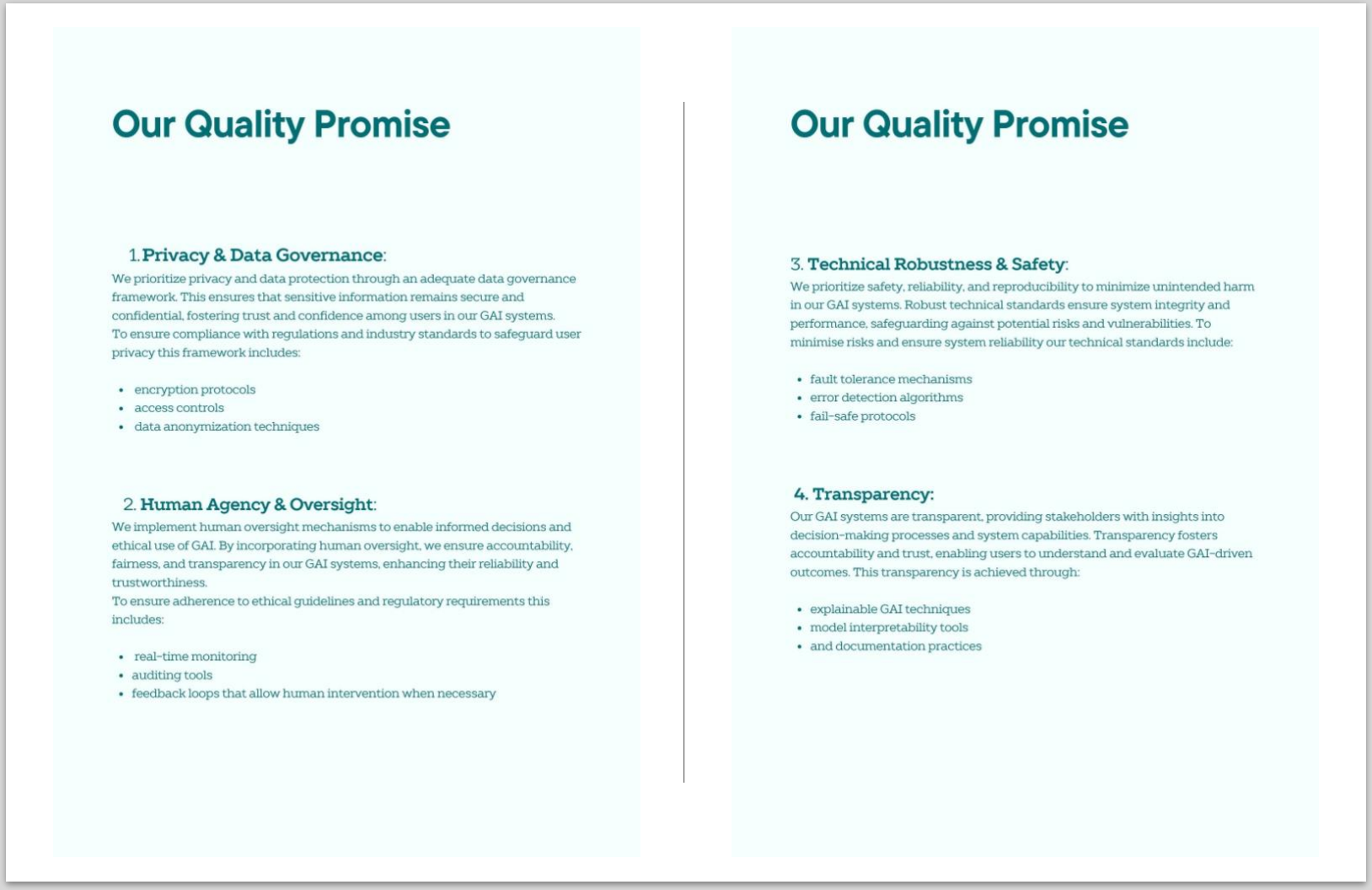


Figure 10 *– Stimuli Testgroup 2, moderate MR*



Our Quality Promise

1. Privacy & Data Governance:

We prioritize privacy and data protection through an adequate data governance framework. This ensures that sensitive information remains secure and confidential, fostering trust and confidence among users in our GAI systems.
To ensure compliance with regulations and industry standards to safeguard user privacy this framework includes:

- encryption protocols
- access controls
- data anonymization techniques

2. Human Agency & Oversight:

We implement human oversight mechanisms to enable informed decisions and ethical use of GAI. By incorporating human oversight, we ensure accountability, fairness, and transparency in our GAI systems, enhancing their reliability and trustworthiness.
To ensure adherence to ethical guidelines and regulatory requirements this includes:

- real-time monitoring
- auditing tools
- feedback loops that allow human intervention when necessary

Our Quality Promise

3. Technical Robustness & Safety:

We prioritize safety, reliability, and reproducibility to minimize unintended harm in our GAI systems. Robust technical standards ensure system integrity and performance, safeguarding against potential risks and vulnerabilities. To minimise risks and ensure system reliability our technical standards include:

- fault tolerance mechanisms
- error detection algorithms
- fail-safe protocols

4. Transparency:

Our GAI systems are transparent, providing stakeholders with insights into decision-making processes and system capabilities. Transparency fosters accountability and trust, enabling users to understand and evaluate GAI-driven outcomes. This transparency is achieved through:

- explainable GAI techniques
- model interpretability tools
- and documentation practices

Figure 10 – Stimuli Testgroup 2, moderate MR

Our Quality Promise

5. **Diversity, Non-discrimination, Fairness**:

We promote diversity and fairness in our AI systems to ensure accessibility and mitigate unfair biases. Embracing diversity enhances inclusivity and fairness, making our GAI technologies more equitable and impactful for all users. These efforts aim to mitigate biases, ensure equal treatment, and promote inclusivity in our GAI technologies – we ensure this through:

- fairness-aware learning algorithms
- diversity-aware training data selection

6. **Societal & Environmental Well-Being**:

We consider the social impact and environmental consequences of our GAI systems, prioritizing the well-being of individuals and communities. By addressing societal and environmental concerns, we aim to create GAI solutions that positively contribute to society. To minimize adverse effects and maximize positive contributions to society and environment this includes:

- conducting impact assessments
- stakeholder consultations
- sustainability analyses

Our Quality Promise

7. **Accountability**:

We uphold accountability in our GAI systems by implementing robust mechanisms for responsibility and oversight. Accountability ensures that stakeholders are held accountable for GAI outcomes, promoting trust and integrity in our operations. To ensure transparency and traceability our technical infrastructure include:

- blockchain-based accountability frameworks
- immutable logs

X

Figure 11 – *Stimuli Testgroup 3, high MR*

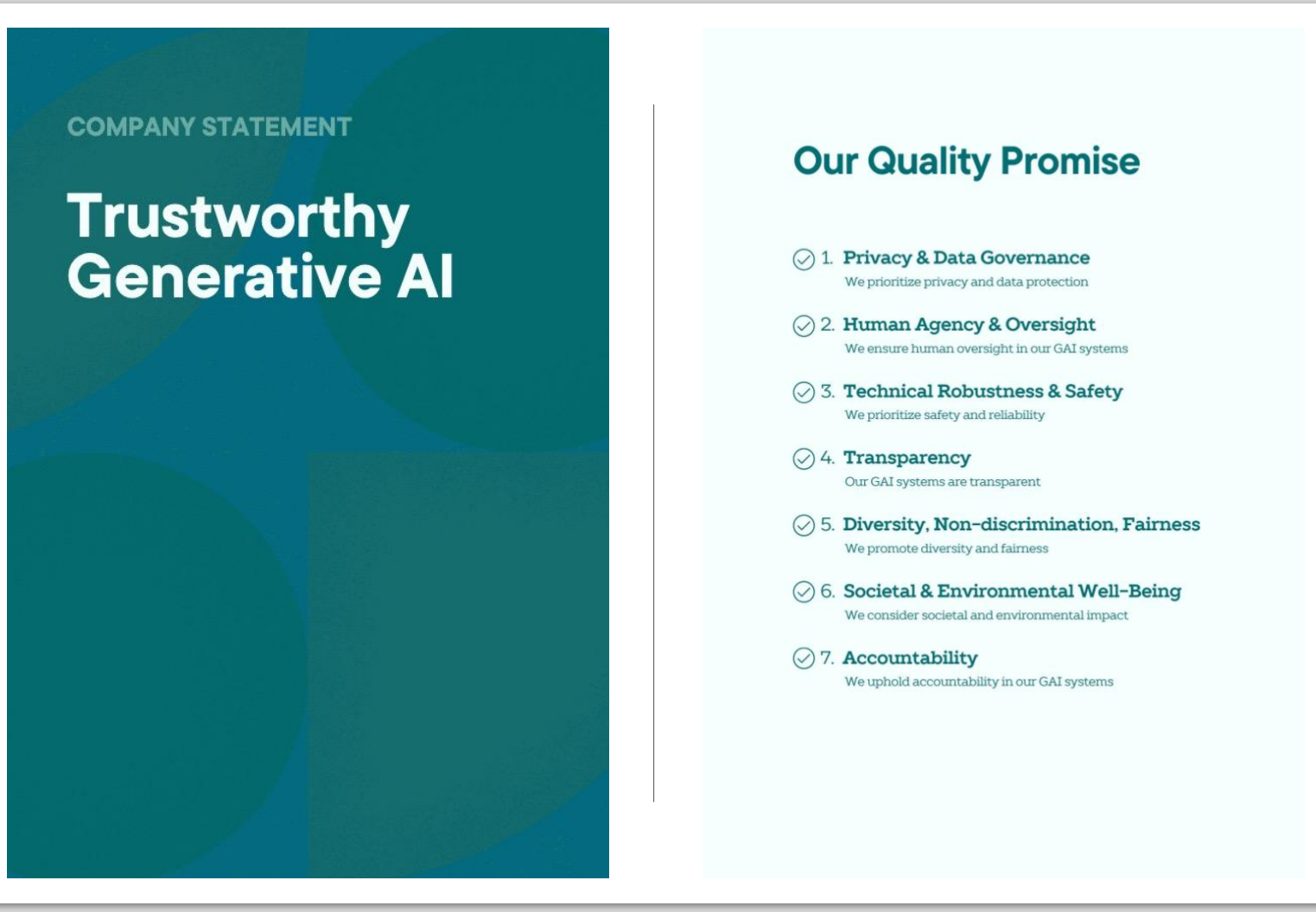


Figure 11 *– Stimuli Testgroup 3, high MR*

Our Quality Promise

1. **Privacy & Data Governance**:

We prioritize privacy and data protection through an adequate data governance framework. This ensures that sensitive information remains secure and confidential, fostering trust and confidence among users in our GAI systems.

This involves implementing encryption protocols such as AES (Advanced Encryption Standard) with 256-bit keys for data at rest and TLS (Transport Layer Security) for data in transit. We utilize cryptographic hashing algorithms like SHA-256 (Secure Hash Algorithm) to ensure data integrity. Additionally, we enforce access controls using role-based access control (RBAC) and attribute-based access control (ABAC) mechanisms, alongside strong authentication methods like multi-factor authentication (MFA) to prevent unauthorized access. Data anonymization techniques including differential privacy and homomorphic encryption are employed to preserve privacy while allowing for meaningful analysis

Our Quality Promise

2. **Human Agency & Oversight**

We implement human oversight mechanisms to enable informed decisions and ethical use of AI. By incorporating human oversight, we ensure accountability, fairness, and transparency in our GAI systems, enhancing their reliability and trustworthiness.

To ensure adherence to ethical guidelines and regulatory requirements this includes implementing robust human oversight mechanisms by integrating real-time monitoring systems powered by anomaly detection algorithms such as Isolation Forests and One-Class SVMs (Support Vector Machines). These systems continuously monitor model performance and trigger alerts for deviations from expected behavior. Human-in-the-loop approaches utilize active learning strategies and uncertainty sampling methods to identify edge cases and ambiguous instances that require human intervention. Version control systems and distributed ledger technologies like blockchain ensure traceability and auditability of model updates and decisions. Automated governance frameworks powered by smart contracts enforce ethical guidelines and regulatory compliance.

Figure 11 *– Stimuli Testgroup 3, high MR*

Our Quality Promise

3. **Technical Robustness & Safety**:

We prioritize safety, reliability, and reproducibility to minimize unintended harm in our GAI systems. Robust technical standards ensure system integrity and performance, safeguarding against potential risks and vulnerabilities.

We prioritize technical robustness and safety through a multi-layered approach that includes fault tolerance mechanisms, error detection algorithms, and model validation techniques. Fault tolerance is achieved through redundancy and failover mechanisms such as active-passive clustering and load balancing. Error detection algorithms including statistical process control and Bayesian change point detection identify anomalies and drifts in data and model behavior. Model validation involves rigorous testing methodologies such as stress testing, fuzz testing, and chaos engineering to evaluate system resilience and performance under adverse conditions. Continuous integration and continuous deployment (CI/CD) pipelines automate testing and deployment processes, ensuring rapid response to emerging threats and vulnerabilities.

Our Quality Promise

4. **Transparency**:

Our GAI systems are transparent, providing stakeholders with insights into decision-making processes and system capabilities. Transparency fosters accountability and trust, enabling users to understand and evaluate GAI-driven outcomes.

Our AI systems prioritize transparency through model interpretability techniques such as SHAP (SHapley Additive exPlanations) values and LIME (Local Interpretable Model-Agnostic Explanations). These techniques generate feature importance scores and highlight the contribution of individual features to model predictions, enabling stakeholders to understand model behavior and identify potential biases. Explanation generation models based on neural attention mechanisms provide contextually relevant explanations for model decisions. Model versioning and lineage tracking using distributed ledger technologies ensure traceability and accountability of model predictions over time. Automated documentation frameworks leverage natural language processing (NLP) and knowledge graphs to generate comprehensive documentation of model architecture, data pipelines, and performance metrics.

Figure 11 – *Stimuli Testgroup 3, high MR*

Our Quality Promise

5. **Diversity, Non-discrimination, Fairness**:

We promote diversity and fairness in our GAI systems to ensure accessibility and mitigate unfair biases. Embracing diversity enhances inclusivity and fairness, making our GAI technologies more equitable and impactful for all users. These efforts aim to mitigate biases, ensure equal treatment, and promote inclusivity in our GAI technologies.

We promote diversity, non-discrimination, and fairness in our GAI systems through algorithmic fairness measures and bias mitigation strategies. Fairness-aware learning algorithms utilize causal inference techniques and counterfactual reasoning to identify and mitigate biases in training data and model predictions. Bias detection algorithms including adversarial attacks and membership inference attacks analyze model outputs and identify potential vulnerabilities. Synthetic data generation techniques such as generative adversarial networks (GANs) and variational autoencoders (VAEs) augment training datasets and address data scarcity issues, ensuring equitable representation across diverse demographic groups. Active learning strategies and adaptive sampling methods mitigate label imbalance and sampling bias, improving model performance and fairness.

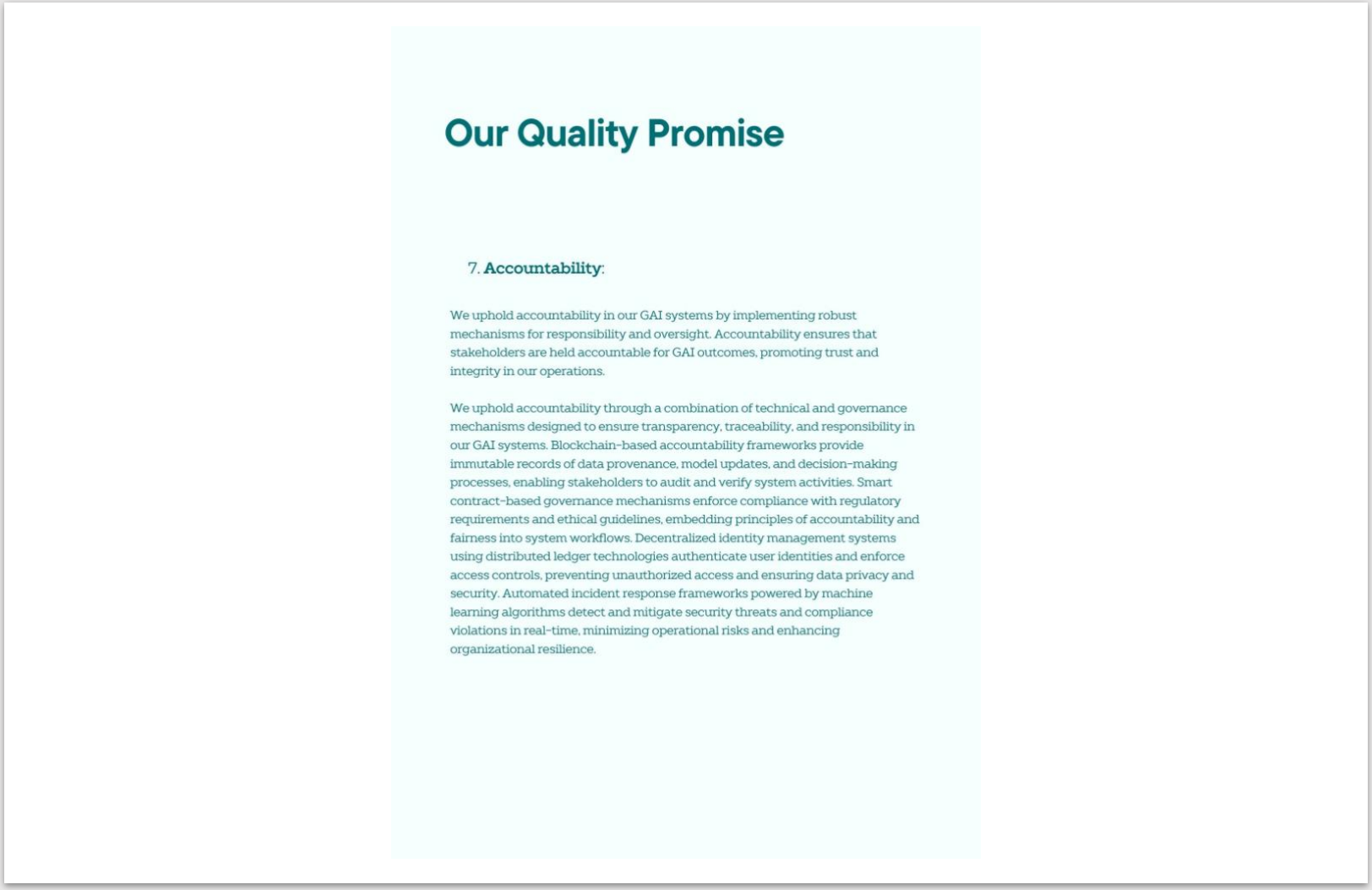
Our Quality Promise

6. **Societal & Environmental Well-Being**:

We consider the social impact and environmental consequences of our GAI systems, prioritizing the well-being of individuals and communities. By addressing societal and environmental concerns, we aim to create GAI solutions that positively contribute to society. To minimize adverse effects and maximize positive contributions to society and environment this includes:

We prioritize societal and environmental well-being in our GAI initiatives by integrating socio-technical impact assessments and environmental impact assessments into our development lifecycle. Socio-technical impact assessments evaluate the potential social, economic, and ethical implications of GAI applications, incorporating stakeholder perspectives and community feedback. Environmental impact assessments quantify the carbon footprint, energy consumption, and resource utilization associated with AI infrastructure and operations, informing sustainable design choices and eco-friendly practices. Decentralized GAI frameworks powered by federated learning and edge computing minimize data transfer and processing overhead, reducing environmental impact and enhancing data privacy and security.

Figure 11 *– Stimuli Testgroup 3, high MR*



Our Quality Promise

7. **Accountability**:

We uphold accountability in our GAI systems by implementing robust mechanisms for responsibility and oversight. Accountability ensures that stakeholders are held accountable for GAI outcomes, promoting trust and integrity in our operations.

We uphold accountability through a combination of technical and governance mechanisms designed to ensure transparency, traceability, and responsibility in our GAI systems. Blockchain-based accountability frameworks provide immutable records of data provenance, model updates, and decision-making processes, enabling stakeholders to audit and verify system activities. Smart contract-based governance mechanisms enforce compliance with regulatory requirements and ethical guidelines, embedding principles of accountability and fairness into system workflows. Decentralized identity management systems using distributed ledger technologies authenticate user identities and enforce access controls, preventing unauthorized access and ensuring data privacy and security. Automated incident response frameworks powered by machine learning algorithms detect and mitigate security threats and compliance violations in real-time, minimizing operational risks and enhancing organizational resilience.