CAMPUS 02 GRAZ
FACHHOCHSCHULE DER WIRTSCHAFT

Consumer perception: How creative is ChatGPT really?

A quantitative questionnaire

Master Thesis

Submitted by: **Okan Karakas**

Student-ID: **01552510**

Program: **Master program in Digital Marketing Management**

Supervisor: **Dr Dietmar Kappel**

Graz, 20.07.2023

I hereby affirm that this research work was written solely by myself and that I have not previously submitted this work at another educational institution for the purpose of receiving an academic degree. Sentences or parts of sentences quoted are marked as such and have been appropriately cited. I further affirm that I have not used other than the declared sources. The submitted electronic version of the thesis matches the printed version.

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benützt und die benutzten Quellen wörtlich zitiert sowie inhaltlich entnommene Stellen als solche kenntlich gemacht habe. Die Arbeit oder Teile der Arbeit wurde(n) bisher in gleicher oder ähnlicher Form weder dieser noch einer anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Die vorliegende Fassung entspricht der eingereichten elektronischen Version.

Graz, 20.07.2023 __________________________________

Okan Karakas

Abstract

Consumer perception: How creative is ChatGPT really?

Okan, KARAKAS

English Abstract:

This master thesis discusses AI technologies and applications such as ChatGPT and their potential implications for professional copywriters. More specifically, this study seeks to clarify whether ChatGPT is a threat to the copywriter's profession or merely serves as a helpful tool in the creative process. The literature review carried out indicates the importance of creativity in advertising and its potential role in influencing consumer perception and behaviour. Thus, the research goal addressed in this study is to compare consumer perceptions of AI-generated and human-written advertising copy in terms of creativity. Based on prior research, it was expected that human intervention has a positive impact on the evaluation of (creative) output (H1). Furthermore, a positive correlation was assumed between attitudes towards (creative) AI and the evaluation of AI-generated advertising copy (H2). For the purpose of testing these two hypotheses, an online questionnaire survey with a sample of 75 participants was conducted. The results suggest that the first hypothesis cannot be fully supported by the data. Nevertheless, significant differences were found in the evaluation between the two advertising copies. As for the second hypothesis, there seems to be no positive correlation between attitude towards (creative AI) and the evaluation of AI-generated advertising copy. With the EU AI Act still being drafted and some researchers calling for AI-generated content to be labelled, the findings of this study hold scientific and managerial significance by revealing consumer perceptions of AI-generated copy and providing insights for advertising professionals.

Keywords: ChatGPT, Artificial Intelligence, Marketing, Copywriting, Consumer Perception, Advertising Creativity

German Abstract:

Diese Masterarbeit befasst sich mit KI-Technologien sowie deren Applikationen wie ChatGPT und ihren möglichen Auswirkungen auf professionelle Werbetexter:innen. Konkret soll diese Studie klären, ob ChatGPT eine Bedrohung für den Beruf des professionellen Werbetexters bzw. der professionellen Werbetexterin darstellt oder lediglich als hilfreiches Werkzeug im kreativen Prozess dient. Die durchgeführte Literaturrecherche verdeutlicht die Bedeutung von Kreativität in der Werbung und ihre potenzielle Rolle bei der Beeinflussung der Kundenwahrnehmung und des Kundenverhaltens. Das Forschungsziel dieser Studie besteht daher darin, die Kundenwahrnehmung von KI-generierten und von Menschen geschriebenen Werbetexten in Bezug auf Kreativität zu vergleichen. Auf der Grundlage früherer Untersuchungen wurde erwartet, dass die menschliche Intervention einen positiven Einfluss auf die Bewertung der (kreativen) Texte hat (H1). Darüber hinaus wurde ein positiver Zusammenhang zwischen der Einstellung gegenüber (kreativer) KI und der Bewertung von KI-generierten Werbetexten angenommen (H2). Zur Überprüfung dieser beiden Hypothesen wurde eine Online-Fragebogenerhebung mit einer Stichprobe von 75 Teilnehmer:innen durchgeführt. Die Ergebnisse deuten darauf hin, dass die erste Hypothese durch die Daten nicht vollständig gestützt werden kann. Dennoch wurden signifikante Unterschiede bei der Bewertung der beiden Werbetexte festgestellt. Was die zweite Hypothese betrifft, so scheint es keinen positiven Zusammenhang zwischen der Einstellung gegenüber (kreativer) KI und der Bewertung von KI-generierten Werbetexten zu geben. Angesichts der Tatsache, dass der EU AI Act noch im Entwurf steht und einige Forscher:innen eine Kennzeichnungspflicht für KI-generierte Inhalte fordern, sind die Ergebnisse dieser Studie von wissenschaftlicher und wirtschaftlicher Bedeutung, da sie Aufschluss über die Kundenwahrnehmung von KI-generierten Werbetexten geben und Erkenntnisse für Werbefachleute liefern.

Schlüsselwörter: ChatGPT, Künstliche Intelligenz, Marketing, Werbetexten, Kundenwahrnehmung, Werbekreativität

Table of Contents

List of tables

List of figures

Introduction

The conceptualisation of the Turing Test is widely regarded as the seminal event that marked the emergence of chatbot technology (Adamopoulou & Moussiades, 2020, p. 2). Originally formulated by Alan Turing (Turing, 1950), the question of whether it is plausible for machines to converse with humans while successfully hiding their machine nature from the human participants has evolved into a much more complicated question in light of advances in Artificial Intelligence (hereinafter, AI). As a result of remarkable advances in Deep Learning (hereinafter, DL), a specific subset of AI, chatbots have achieved an expanded repertoire of capabilities that goes far beyond simply being able to participate in everyday conversations (Vogt, 2018, p. 690).

One chatbot, the Chatbot Generative Pre-trained Transformer (hereinafter, ChatGPT), stands out in particular because of its capabilities and potential areas of use. In fact, the novel chatbot (and AI in general) has been experiencing a lot of “hype” in the media and broader society since its launch in November 2022 (ChatGPT & Affairs, 2022, p. 379; Salvagno et al., 2023, p. 1; Tlili et al., 2023, pp. 1–2). Unlike conventional chatbots, ChatGPT is not only able to conduct natural conversations on a wide range of topics, but can also be used to produce creative texts, among other capabilities (Darlington, 2023, p. 50; Taecharungroj, 2023, p. 5; Zhou et al., 2023, p. 2).

While some researchers report predominantly on the benefits of this AI evolution, such as increased work efficiency (Volkmar et al., 2021, p. 367; Qin & Jiang, 2019, p. 338), other researchers are more critical of the development, also with regard to job security. Some researchers are arguing that AI and applications such as ChatGPT will soon not only assist in various professions but will make these jobs obsolete altogether (Serdouk & Bessam, 2023,

p. 104). This newly emerging AI trend could also affect the job of professional copywriters. To date, the question remains open as to whether AI will complement or completely replace creative professions such as that of the professional copywriter. Due to ChatGPT's recent nature, scientific evidence on AI's impact on the profession of copywriters is scarce. While there are plenty of studies dedicated to chatbots, previous research mainly focused on rather technical aspects (Taecharungroj, 2023, pp. 1–3).

For this reason, this study focuses on AI (i.e., ChatGPT) and the impact of these technologies and applications on the copywriter's job. More specifically, customer perceptions of AI-generated versus human-written advertising copy in terms of creativity will be investigated through an online questionnaire survey. In addition, the impact of attitudes towards AI (and its potential ability to be creative) on customer perceptions will be explored.

In order to ensure a systematic structure of this paper, the following chapter organization is employed: Chapter 1 of this paper provides a comprehensive literature review that delves into the broad field of AI, chatbots, specifically focusing on ChatGPT, and the usage of AI technology in the field of marketing. The literature review begins by exploring the major subfields of AI. It covers significant AI fields and technologies relevant for ChatGPT, from early AI approaches (e.g., expert systems) to the emergence of machine learning and deep learning techniques that have revolutionized the field. The review then moves on to the design dimensions of chatbots and the growing prominence of chatbots in various industries, including customer service and marketing. A substantial part of the literature review is devoted to a comprehensive examination of ChatGPT. It covers its underlying architecture, which is based on transformer models, and how it differs from earlier versions of so-called language models. This chapter also discusses the ChatGPT training process, including pre-training and fine-tuning, as well as the limitations and challenges associated with large-scale language models, such as ethical concerns and potential bias. The literature review then focuses on the usage of AI technology in the field of marketing. It explores the different use cases of AI in marketing and examines the state of research in various marketing AI fields and reveals a research gap in that domain. The report highlights how AI has transformed marketing strategies, enabling companies to improve customer experiences, optimise marketing campaigns and increase overall efficiency and effectiveness.

Building on the findings of the literature review, Chapter 2 presents the theoretical framework underlying the research question presented in Chapter 1. It forms the conceptual basis for the study of the impact of ChatGPT on the profession of professional copywriting. The chapter outlines the hypotheses to be tested in a quantitative analysis. Hypothesis 1 investigates the differences between AI-generated and human-written copywriting in terms of creativity. Hypothesis 2 examines the possible influence of attitudes towards creative AI on the evaluation of AI-generated advertising copy.

Chapter 3 explains the applied research methodology in detail. This includes the research design, data collection methods and the specific measures used to assess the perceived advertising creativity of the written advertising texts. This chapter explains the sample selection process and the rationale for using a convenience sample, highlighting its limitations and potential impact on the generalisability of the findings.

Chapter 4 presents the findings and results from the analysis of the collected data. It provides a detailed account of the differences between AI-generated and human-written copy in terms of perceived advertising creativity i.e., it’s underlying construct. The chapter analyses the statistical significance of these differences and explores the implications for the profession of professional copywriting. Furthermore, the relationship between the attitude towards (creative) AI and the evaluation of AI-generated advertising copy is examined. It also justifies the choice of data analysis methods used to test the hypotheses and draw meaningful conclusions from the data collected.

Chapter 5 summarises and discusses notable observations from the study. The chapter draws final conclusions based on the findings and discusses their implications for both the advertising industry, the job of the professional copywriter, and the broader field of AI application in marketing. It also reflects on the contribution of the study to the existing literature.

The final chapter, Chapter 6, addresses the limitations encountered during the research process. It acknowledges the limitations that may have influenced the findings of the study and suggests ways to address these limitations. In addition, Chapter 6 identifies possible avenues for further research, highlighting areas that were not fully explored in this study. It highlights the value of exploring additional text formats, different product categories, cross-linguistic effects and integrating qualitative data to gain more comprehensive insights into the impact of ChatGPT and creative AI applications in marketing.

Chapter 1: Prior Literature

Although AI may seem like an innovation to many non-experts, the technology has basically been around since the birth of the computer (Sumitha, 2022, p.18). More specifically, the literature indicates that AI originated in 1956 at the Dartmouth Summer Conference, the first held AI conference (Volkmar et al., 2021, p. 360).

Despite years of existence and advancement of AI, researchers still are divided on a unified definition. The table below shows a selection of AI definitions from various researchers in recent years. As can be seen in the table, researchers have taken slightly different approaches to define AI over the years.

Table 1 - AI definitions[1]

Consec. number	Citation	Definition
1	(Volkmar et al., 2021, p. 361)	Artificial Intelligence, as a discipline of science and technology, is capable of completing many tasks intelligently, identifying faults and learning from them. Thus, AI has the potential to acquire intelligent behaviour and operate correctly in an unpredictable environment.
2	(Bünte, 2021, p. 32)	Any system that utilizes algorithms to acquire knowledge - either with or without human guidance.
3	(Qin & Jiang, 2019, p. 338)	*"[…], a set of disruptive technologies which simulate human intelligence and realize machine intelligence, […]"*
4	(Baumgarth & Kirkby, 2022, p. 33)	All technical attempts to emulate human intelligence fall under the umbrella of Artificial Intelligence.
5	(Broussard et al., 2019, p. 673)	*"[…] a branch of computer science focused on simulating human intelligence"*

[1] Definitions 1, 2 & 4: Paraphrased German-English translations by the author of this work according to APA citation rules.

Table 1 - Continued

6	(Castro & New, 2016, p. 2)	*“AI is a field of computer science devoted to creating computing machines and systems that perform operations analogous to human learning and decision-making”*
7	(Kaplan & Haenlein, 2019, p. 17)	*“A system’s ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation”*
8	(Shankar, 2018, p. vi)	*“Programs, algorithms, systems or machines that demonstrate intelligence. More generally, it is used to denote a set of tools that can enhance the intelligence of a product, service, or solution”*
9	(Serdouk & Bessam, 2023, p. 101)	*“This technology is known as artificial intelligence (AI) because it simulates the human mind in its intelligence through programmes and applications that help the institution perform some of its activities with accurate effectiveness that human efforts may not reach. On the other hand, it contributes to ideal decision-making, guarantees the achievement of the institution’s current goals, and draws its long-term development plans”*
10	(McCarthy, 1998, p. 2)	*“It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable”*
11	(Russell & Norvig, 2021, p. 19)	*“The field of artificial intelligence, or AI, is concerned with not just understanding but also building intelligent entities—machines that can compute how to act and safely in a wide variety of novel situations”*

Note. This table was created by the author of this paper.

Nevertheless, the different definitions have some similar aspects. Volkmar et al. group those similarities into the categories “technological focus” and “human focus”. The technological focus relies on the idea that AI is capable of thought, action, perception of the environment, and independently completing challenging tasks. On the other hand, the human focus highlights that AI must demonstrate a degree of intelligence comparable to that anticipated by humans when carrying out a certain task (Volkmar et al., 2021, p. 360). Based on these two key features, in this paper AI is considered to be any technology that is capable of emulating human intelligence.

Furthermore, AI can be distinguished based on performance or “strength”. This distinction is made between weak AI (Artificial Narrow Intelligence) and strong AI (Artificial General Intelligence; Abbr.: AGI). While weak AI can only replace specific human characteristics i.e., carry out certain tasks, strong AI (i.e., AGI) integrates machine and deep learning for processing and executing specific tasks. By doing so, AGI finds relationships and analogies in vast quantities of data through algorithms to perform complicated tasks beyond the scope of human abilities. This transferrable "intelligence" enables the handling of new obstacles. So, while AGI evolves independently and adapts to changing contexts, weak AI restricts itself to a specific set of tasks. Thus, the adaptive capability of AGI makes it a difficult endeavour for AI researchers, which is why most AI systems i.e., applications today are based on weak AI (Baumgarth & Kirkby, 2022, p. 33; Hauck & Pagel, 2020, p. 51), and AGI is still in the research and development phase.

AI Subfields

The following subsections describe the major subfields of AI. This categorisation serves the reader to identify the technologies behind ChatGPT. It should be noted, however, that these subfields frequently overlap and support each other (de-Lima-Santos & Ceron, 2021, p.

16). The categorization was carried out by the authors de-Lima-Santos & Ceron, who divided the different AI applications into the subfields of **(1)** Machine Learning, **(2)** Computer Vision, **(3)** Speech Recognition, **(4)** Robotics, **(5)** Planning, Scheduling, and Optimization, **(6)** Expert Systems, and **(7)** Natural Language Processing (de-Lima-Santos & Ceron, 2021, p. 13). It is worth mentioning that additional instances (e.g., Deep Learning as an instance of Machine Learning) can be assigned to the various AI subfields. However, those instances are intentionally omitted from the scope of this work.

Machine Learning (ML)

Machine Learning (hereinafter, ML) entails the analysis of large databases. By interpreting the data, a system is provided with, the algorithms used in ML can find trends i.e., patterns and provide predictions and suggestions for its users. Furthermore, as they acquire access to new data, these algorithms can enhance the accuracy and/or the output they provide. The underlying concept of ML software is based on the use of algorithms as well as statistical models to acquire knowledge and adapt to new data over time. This process encompasses multiple runs of the ML software through a continually growing database, where the application actively changes the parameter values of an AI model throughout the data analysis, either with the help of an AI expert or by self-regulation. This kind of "learning" can be accomplished by four different methods: Supervised, unsupervised, semi-supervised, and reinforcement learning (Kursh, 2021, pp. 14-23).

Supervised machine learning is an approach that establishes an association between input values and their associated output values by using classified examples of input-output pairs. In contrast, unsupervised machine learning does not use a labelled dataset. In this method, patterns in the data are revealed by exposing the underlying data structure, e.g., the grouping of related components. As the name suggests, semi-supervised learning uses a combination of both learning approaches mentioned previously, so it uses labelled and unlabelled data to enhance its capabilities. Lastly, there is the reinforcement learning (RL) method. The basic idea behind reinforcement learning is to examine how AI agents make decisions in a particular setting and determine the most effective actions to take in order to optimize their total reward (Samoili et al., 2021, p. 14).

As the focus of this study is on ChatGPT, it should be noted that the AI application relies heavily on the use of ML technology (Hacker, 2023, p. 2).

Computer Vision (CV)

A general definition of Computer Vision (hereinafter, CV) is a technology that analyses and decodes digital images to recognize and categorize objects, people, and other visual elements in the provided data (Samoili et al., 2021, p. 25). To be more precise, CV researchers apply mathematical and statistical methods to recover the 3D model of an object in an image. Nowadays, CV is already an integral part of our (professional) everyday life. For example, the technology is used in unlocking smartphones (fingerprint recognition and face detection), autonomous driving, warehouse logistics (autonomous package delivery) and machine inspection in production plants. Nevertheless, the above-mentioned application areas are specialized for certain image material and refer to rather narrow functions. This is partly because CV is an inverse problem, where an attempt is made to determine a variable for which there is insufficient information. CV researchers resort to physical and probabilistic models or ML to determine possible solutions (Szeliski, 2022, pp. 3-8).

(Automatic) Speech Recognition (ASR)

Amazon's Alexa, Apple's Siri, or modern navigation systems are just a few examples from our daily lives that rely on the technology behind speech recognition, also known as automatic speech recognition (hereinafter, ASR). Today, one-third of human-machine interaction already takes the form of voice rather than keyboard commands (Russell & Norvig, 2021, p. 47). Converting spoken sound into text is a simple definition of ASR (Russell & Norvig, 2021, p. 900).

Robotics

Robotics is an AI subfield that combines various neural technologies in order to enable AI-powered machines and systems to perform different functions and tasks in collaboration with humans in uncertain settings i.e., environments (de-Lima-Santos & Ceron, 2021, p. 16). To be able to act and make decisions in these environments, robots (i.e., physical agents) are equipped with effectors e.g., with legs and wheels. When the effectors are called into action, physical force is applied to the environment, which can have three consequences: the state of the robot changes, and/or the state of the environment changes, and/or the state of the people in the environment changes. To perceive these state changes, robots are additionally equipped with numerous sensors. Cameras and lasers, for example, measure the state of the surroundings and people around the agent, while gyroscopes and accelerometers measure the state of the robot itself. These robots can range from simple machines in assembling lines to fully autonomous legged robots like the ones we know from science fiction movies (Russell & Norvig, 2021, pp. 932-933).

Planning, Scheduling, and Optimization

The AI subfield "Planning, Scheduling and Optimization" is an area that already has a higher degree of maturity (de-Lima-Santos & Ceron, 2021, p. 16). AI Planning, Scheduling, and Optimization solves common problems of everyday (business) life, where resources need to be optimally allocated to achieve intended results. AI Planning here can be seen as a sequence of specific actions to accomplish defined goals. Scheduling and Optimization on the other hand have the task of identifying the optimal sequence of actions to achieve the defined goals with optimised use of resources. Application examples for Planning, Scheduling, and Optimization are logistics and transport systems, which rely on constant improvements in an international competition environment (Salido, 2010, pp. 1-2).

Expert Systems (ES)

One of the early AI subfields, expert systems (also known as “knowledge-based systems”), gained its momentum in the 1970s. The task of such a knowledge-based system is to reach or even surpass the expertise level of a specialist in a certain domain based on given data (Russell & Norvig, 2021, p. 356). The initial goal of this technology was to enable knowledge transfer at minimal cost and to become less dependent on expert opinions. Facts, rules, heuristics, and procedures are provided to the expert system on the grounds on which it makes decisions (Leo Kumar, 2019, p. 4767).

Natural Language Processing (NLP)

Natural Language Processing (hereinafter, NLP) was invented to simplify the interaction between humans and machines, since many people do not have programming skills or know machine-specific languages (Pais et al., 2022, p. 3). Moreover, NLP's goal is to use AI tools in combination with linguistics, cognitive psychology, and neurobiology, to improve empirical knowledge of languages and language use by combining linguistics, cognitive psychology, and neuroscience (Russell & Norvig, 2021, p. 874).

This rather broad AI subfield includes use cases such as translations, summaries, content analysis, dialogue generation, and image/video captioning (de-Lima-Santos & Ceron, 2021, p.15; Veel, 2018, p.3; Sai et al., 2022, p. 2). NLP can be defined as a technology that can understand and create human language (Skibba, 2020, p. 723).

Furthermore, NLP includes two additional instances, namely Natural Language Understanding (hereinafter, NLU) and Natural Language Generation (hereinafter, NLG) (Dong et al., 2023, p. 2; Gatt & Krahmer, 2018, p. 68). As the central focus of this work is on the creation of AI-generated copywriting, these instances will be addressed in more detail in the remainder of this section.

Natural Language Understanding & Generation

As mentioned above, NLP consists of additional instances, namely NLU and NLG. The NLU instance is further divided into two parts: intent recognition and slot filling. The goal of intent recognition is to anticipate the user's purpose i.e., intent, whereas the goal of slot filling is to identify connected entities within a database. The technique of NLU is usually being used to analyse many documents or texts in a fast fashion. Application examples for NLU can be found in content analysis (e.g., sentiment analysis) and text categorization. During the NLU process, unstructured data (e.g., natural language) is analysed and consequently transformed into structured data (i.e., formal representations) (Chuang & Cheng, 2022, pp. 1–5). On the other hand, the goal of NLG is to create texts and, when using chatbots such as ChatGPT, to provide answers to questions and ask follow-up questions when needed. Dong et al., f. ex., define NLG as *"the process of producing a natural language text to meet specified communicative goals"* (Dong et al., 2023, p. 2). Educating a model that accepts data as input, understands the context from that input, and produces original text that is pertinent to the input area are crucial steps of the NLU and NLG instances. The final text i.e., output must successfully communicate the intended message while adhering to the fundamental principles of language structure (Fatima et al., 2022, p. 53490). Some examples of the use of NLG are translations, summaries, or dialogue generation (Sai et al., 2022, p. 2). Furthermore, natural language can be generated at different levels, including character, word, and phrase level. Sentence-level generation thoroughly examines the text and comprehends sentence-context connections. The next word in a series is predicted by word-based creation. Individual figures are identified through character-level identification i.e., creation (Fatima et al., 2022, p. 53490).

Finally, NLG can also be distinguished between text-to-text generation and data-to-text generation (Gatt & Krahmer, 2018, p. 68). As this distinction is crucial for a precise framing of this paper, it will be discussed in the following.

Data-to-Text Generation

As defined by Gatt and Krahmer, data-to-text generation is the process of automatically generating text from non-linguistic data. In the process of this technique, information is extracted from various sources such as databases or spreadsheets and converted into natural language text (Gatt & Krahmer, 2018, p. 66).

This type of text generation has become particularly popular with news publishers in recent years. Also known under the umbrella term "algorithmic journalism", "computational journalism", "robotic journalism", "automated journalism" and other similar terms describing the same concept (Kotenidis & Veglis, 2021, pp. 244-245) for the whole news production/distribution process, this type of (semi-) automatic news production has been used by various media outlets, such as the New York Times, for mainly data-driven news (e.g., earthquake reports or sports news) (Melin et al., 2018, p. 43356). Many researchers argue that this type of automated news generation may allow journalists to avoid repetitive tasks, save time and use the newly gained time for conducting interviews or writing particularly important news stories (Noain-Sánchez, 2022, p. 112). For further reference regarding “algorithmic journalism” see (de-Lima-Santos & Ceron, 2021; Melin et al., 2018; Noain-Sánchez, 2022; Serdouk & Bessam, 2023). Another intriguing application of data-to-text generation is text generation from pictures, which is especially interesting for image captioning due to the abundance of image data available on the internet (Dong et al., 2023, p. 2).

Text-to-Text Generation

The main difference between data-to-text generation and text-to-text generation can be determined by the input the two instances of NLP use (Gatt & Krahmer, 2018, p. 68). As mentioned above, data-to-text generation is based on non-linguistic data. Text-to-text generation, on the other hand, takes existing text, converts it to semantic representations, and creates the desired output accordingly. This task i.e., model is also known as "sequence-to-sequence" (abbr.: Seq2Seq) (Gatt & Krahmer, 2018, p. 66; Yu et al., 2022, p. 2). Lastly, text-to-text generation can be divided into different types of tasks it can fulfil, namely text abbreviation, text expansion, text rewriting, and reasoning. The text abbreviation task is designed to reduce the amount of information in lengthy texts to a manageable or shorter amount. It frequently incorporates study on text summarization, question generation, and distractor generation. Short text expansion and topic-to-essay generation are typical applications of the text expansion task. Here, the goal is to turn the input words into grammatically correct outputs, resulting in full sentences by combining elements such as conjunctions or prepositions. The text rewriting and reasoning task, on the other hand, consists of two additional sub-tasks. While text rewriting takes an already existing text and transfers it to a specific style (e.g., from formal to academic writing), the task of reasoning is to create a whole dialogue and answer to multiple questions in a reasonable manner (Dong et al., 2023, pp. 2–19; Yu et al., 2022, p. 2).

To the best of the author's knowledge, no academic sources were identified in the existing research to confirm ChatGPT's data-to-text generation and/or text-to-text generation capabilities. However, if one asks the chatbot itself, the following answer is provided:

> *"As an AI language model, I am proficient in both text-to-text generation and data-to-text generation tasks. […]" (*ChatGPT*, 2023).*

Chatbots

As early as 1950, Alan Turing asked himself the question of whether a computer program (i.e., machine) is capable of conversing with a group of people without them noticing that a machine is communicating with them. In his article, Turing proposed a thought experiment called the "imitation game" (also known as the Turing Test). In this game, a human judge (called the "interrogator") communicates with another human and a machine through teletype, absent of being able to see or hear each other. The goal of the interrogator is to distinguish between human and machine by asking questions that only a human being might be able to answer. If the machine manages to fool the interrogator into believing it is human, the machine passes the imitation game (i.e., Turing Test) and, according to Turing, demonstrates the ability to think (and converse) like a human. Many regard this game i.e., test as the birth hour of the original chatbot idea (Adamopoulou & Moussiades, 2020, p. 2; Turing, 1950, p. 433).

Generally speaking, a bot can be defined as *"a device or piece of software that can execute commands, reply to messages, or perform routine tasks, as online searches, either automatically or with minimal human intervention (often used in combination)"* (Cai, 2013, p. 116). In recent years, a certain type of bot has become more popular: Algorithm-based dialogue systems i.e., chatbots that make use of AI. These bots, also known in research as conversational agents (Motger et al., 2022, p. 1), chatterbots, conversational interfaces, and similar terms describing the same concept (Motger et al., 2022, p. 7), aim to simulate human-like dialogues by using NLP techniques (i.e., NLU) to understand keywords, phrases, or input in general, and give adequate answers (NLG). This system or software can subsequently be built into websites or apps, for example, using APIs (application programming interface) (Salvagno et al., 2023, p. 1). According to the literature, chatbots are used primarily for productivity reasons. Entertainment, social factors, and interest in innovations are other reasons for using chatbots. Chatbots have long since arrived in the business world, where they can achieve cost savings and serve many users simultaneously round the clock (Adamopoulou & Moussiades, 2020, p.1).

However, not all chatbots are the same and they can be classified/categorized by different design dimensions. Motger et al. describe these larger categories of chatbot differentiation as **(1)** Prescriptiveness, **(2)** Knowledge Base, **(3)** Service, **(4)** Response Generation, **(5)** Interaction, **(6)** Integration, and **(7)** Human Aid (Motger et al., 2022, pp. 16-18).

Design Dimensions

The following subsections provide some explanation of the above mentioned chatbot design dimensions. These design dimensions serve as the basis for classifying ChatGPT according to its distinctive characteristics (see **Classification**).

Prescriptiveness

A key differentiator of chatbots is the goal they pursue. Here one can generally distinguish between task-oriented and non-task-oriented chatbots.

As the name suggests, task-oriented chatbots are designed to achieve very specific goals. These type of chatbots can be found in ticket bookings and order management systems, for example. Task-oriented chatbots are limited to certain areas i.e., domains and cannot answer every type of question in a satisfactory way (Hussain et al., 2019, p. 952). These chatbots are usually integrated into existing software or websites (Motger et al., 2022, p. 16).

In case of textual task-oriented chatbots, most often pattern matching techniques are used to group text and generate contextually suitable responses. This pattern matching feature allows the task-oriented bots to figure out the context (i.e., the user's intent) and respond appropriately and quickly (Hussain et al., 2019, p. 953).

Non-task chatbots, on the other hand, aim to have natural sounding conversations with its users, similar to human-to-human interactions, rather than performing specific tasks such as ticket bookings. They frequently serve to entertain the user as well. These chatbots are divided into two types: Generative-Based and Retrieval-Based models. While Retrieval-Based models mostly rely on a set of ready-made answers, Generative-Based models are able to answer more concrete questions and do not require a corpus of answers, which are limited to a specific domain (Hussain et al., 2019, pp. 953-954). Both, Retrieval-Based and Generative-Based models, use several Artificial Neural networks that transform the user's input into vector representations and generate a response (Adamopoulou & Moussiades, 2020, pp. 7-8).

Knowledge Base

Another important distinctive factor of chatbots is the knowledge base, which determines the extent of the chatbot's conversation capabilities (Motger et al., 2022, p. 16). Generic chatbots can answer questions from all kinds of domains. Quite prominent examples for generic chatbots are Alexa and Google Assistant, which can help users in many different domains i.e., scenarios (Bhirud et al., 2019, p. 225). Chatbots that can converse in more than one domain are called Cross- or Open-Domain chatbots. Lastly, chatbots that rely on a single expert knowledge base, are called Domain-Specific, Domain-Dependent or Closed-Domain chatbots (Adamopoulou & Moussiades, 2020; Motger et al., 2022, p. 17). The goal of Open-

Domain (i.e., Cross-Domain) chatbots is to be able to have conversations that are as human as possible. Closed-Domain chatbots, on the other hand, pay more attention to the accuracy of the answers given, which is very important for educational chatbots, for example (Lin et al., 2023, pp. 5–6).

Service

The service dimension refers to the relationship between the chatbot and the user (Motger et al., 2022, p. 17). Adamopoulou and Moussiades differentiate here between interpersonal, intrapersonal and inter-agent chatbots (Adamopoulou & Moussiades, 2020, p. 3). Interpersonal chatbots do not provide an individual user experience and provide their services based on a broad user classification. Intrapersonal chatbots, on the other hand, tailor their services or responses to the individual needs, and are therefore relying on user profiling (Motger et al., 2022, p. 17). Finally, inter-agent chatbots are worth mentioning. These enable communication between two (or more) different chatbots (Adamopoulou & Moussiades, 2020, p. 3).

Response Generation

Response generation is another design element of chatbots that describes the way appropriate responses are generated based on user input. Essentially, one can distinguish between Deterministic and ML-based mechanisms. Deterministic mechanisms process user input by analysing possible structures and connections and assigning this structured data to a specific response. ML-based mechanisms, on the other hand, use ML (i.e., DL) to process the inputs and generate the answer based on the analysis of large data sets (Motger et al., 2022, p. 17).

Interaction

This design dimension refers to the way a chatbot employs the communication mechanisms to process the input and generate a response (Motger et al., 2022, p. 17; Pérez-Soler et al., 2021, p. 3). Typically, chatbots are text-based, voice-based, or a combination of both (Hussain et al., 2019, p. 947). However, modern chatbots are not limited to text and voice-based input/output. While user input can also take the form of inline buttons, for example, output can include images, videos, links, and other formats (Pereira & Díaz, 2018, p. 4). In addition, further technical developments can have an impact on how these communication mechanisms just mentioned take place. Examples include automatic spelling correction of user input to better understand the user's intent, machine translation to enable the chatbot to have multilingual capabilities, and sentiment analysis to understand the user's current emotional state and take it into account when responding (Adamopoulou & Moussiades, 2020, p. 10).

Integration

The integration dimension is concerned with the environment/ecosystem in which the chatbot has been designed, developed, and deployed, as well as the integration of the chatbot into other software systems and services. With respect to this dimension, Motger et al. conceive a chatbot with different layers/components, namely the interface layer, the dialogue management layer, and the knowledge base layer (Motger et al., 2022, p. 18).

The interface layer describes the area of the chatbot that is visible to the user. Also known as the front end of the dialogue system, this layer acts as the intermediary between the user and the system. The main task of this layer is to collect the user's input and then display the responses generated by the system to the user. To access this interface, users usually have the option of launching a website or a mobile application (standalone app/website) (Bhirud et al., 2019, p. 226). Furthermore, chatbots are often integrated into social media platforms and their messengers, such as the Facebook Messenger. These integrations allow companies to easily reach and interact with a large audience and save on staff costs (Ramesh & Chawla, 2022, p. 477). Developers often must manage (external) data sources on the one hand and third-party software (e.g., through APIs) on the other when integrating chatbots (Pereira & Díaz, 2018, p. 5).

The dialogue management layer has the task of effectively structuring the dialogue between the user and the system, checking the context (the user's intent), and updating it if necessary (Adamopoulou & Moussiades, 2020, p. 10; Pérez-Soler et al., 2021, p. 5). Moreover, the dialogue management layer has other modules that ensure a smooth dialogue, namely the "ambiguity handling", "data handling", and "error handling" modules. The ambiguity handling module helps the chatbot to clarify the user's input (e.g., by asking further questions). The data handling module allows the chatbot to store the user's information in a file and therefore provide a more personalized experience to its users. Lastly, the error handling module assures the proper chatbot operation in unexpected contexts i.e., scenarios (Adamopoulou & Moussiades, 2020, p. 10).

Furthermore, the chatbot accesses knowledge bases (knowledge base layer) in the backend using APIs or databases to serve the users' intent (Adamopoulou & Moussiades, 2020, p. 10). This connection between backend and knowledge base can be made of one or several databases (Zumstein & Hundertmark, 2017, pp. 98-99).

The number of required databases between backend and knowledge base usually depend on the chatbot's goal. Due to the multiplicity of domains it has to handle, Open-Domain chatbots mostly use data that is freely and abundantly available on the internet.

Closed-Domain chatbots, on the other hand, often use data that has been collected specifically for their purpose (Lin et al., 2023, p. 8).

Human Aid

Finally, the design dimension “Human Aid” remains to be mentioned. This dimension describes whether a chatbot can act completely autonomously in the conversation process or with the help of human intervention (Motger et al., 2022, p. 18). The latter is also referred to as human-mediated chatbot (Adamopoulou & Moussiades, 2020, p. 4) or as human-aided bots (Kucherbaev et al., 2018, p. 38).

More specifically, Kucherbaev et al. define human-aided chatbots as conversational agents that need at least one human intervention during runtime (in the “loop”) and therefore exclude chatbots that solely rely on pre-training data (Kucherbaev et al., 2018, p. 37). The authors identify advantages, but also disadvantages of chatbots with human assistance. On the one hand, human intervention provides more robustness and flexibility, as humans can also respond to new input instructions, and therefore decrease uncertainty. On the other hand, with a large number of chatbot users, companies most likely face high personnel costs and cannot provide real-time support due to limited human resources (Kucherbaev et al., 2018, p. 38).

Autonomous chatbots, on the other hand, describe intelligent agents that function independently of human guidance. Self-driving cars, which are already in use, and autonomous weapons systems, which are still in the development stage, are well-known examples of autonomous AI applications i.e., agents. In the field of AI science, however, autonomy describes the degree to which an intelligent agent can achieve a specific goal within a predefined framework. From a philosophical point of view, this kind of "autonomy" does not represent true autonomy if the AI is not able to determine its own goals. The self-driving car, for example, cannot decide for itself to run stop signs (Totschnig, 2020, pp. 1-2).

ChatGPT

ChatGPT is a novel intelligent agent developed by the company OpenAI, which was made available to the public in November 2022 (ChatGPT & Affairs, 2022, p. 379; Salvagno et al., 2023, p. 1; Tlili et al., 2023, p. 2). Designed to have human-like conversations with its users, ChatGPT outperforms many other chatbots with similar capabilities (Darlington, 2023, p. 50; Tlili et al., 2023, p. 1). Not only the performance and capabilities of the new chatbot, but also the social hype around ChatGPT is a novel phenomenon in the world of AI. For instance, just one week after the release of ChatGPT, the CEO of OpenAI, Sam Altman, announced that the 1 million user mark had already been reached (Taecharungroj, 2023, p. 1). One possible explanation for the ChatGPT hype can be found in the AI lab i.e., company OpenAI, already known for innovations, such as DALL·E 2 (*DALL·E 2*, 2023; Zimmerman, 2023, p. 2), which made ChatGPT publicly available as a freemium model (Rana, 2023, p. 7). Another reason for this rapid success may be related to ChatGPT's user-friendly interface (Teubner et al., 2023, p. 96).

However, ChatGPT is not the first chatbot of its kind. ELMo from the Allen Institute for AI, BERT and LaMDA from Google, and RoBERTA from Meta (former Facebook) are just a few well-known examples based on similar technology to ChatGPT (Ge & Lai, 2023, p. 3). So, what exactly makes ChatGPT different from its competitors? The following subsections provide an account of the evolution of the GPT-series, shedding some light on the technological framework behind it, an exploration of the chatbot's capabilities, and a classification based on the design dimensions previously outlined. Subsequently, the limitations of ChatGPT are discussed, along with an assessment of the potential societal and economic consequences resulting from the deployment of this advanced chatbot.

History

OpenAI was founded as a non-profit organization in 2015 and since then has been pursuing the primary goal of being the first AI lab to develop AGI *"—AI systems that are generally smarter than humans—[that] benefits all of humanity"* (Dale, 2021, p. 114; OpenAI, 2023b). With millions of dollars in investments from high-profile individuals such as Elon Musk (Tesla), and Reid Hoffman (LinkedIn), as well as the Microsoft Corporation, OpenAI had a good starting point for achieving this goal (Dale, 2021, p. 114; Lund et al., 2023, p. 5). While the company continues to pursue this goal, OpenAI was restructured in 2019 to become a "capped profit" company consisting of the non-profit OpenAI Foundation and the for-profit OpenAI Limited Partnership (ChatGPT & Affairs, 2022, p. 379). This decision was attributed to the high cost of research and intense AI competition with major industry players like Google. For instance, the estimated cost of GPT-3 (more on the GPT family/series below) per training run amounts to US$ 4.6M (excluding the total development and running costs). Eventually, in the same year, Microsoft agreed to invest $1 billion in OpenAI over the next 10 years, ensuring the liquidity of the company, and received exclusive licensing rights to GPT-3 a year later (Dale, 2021, p. 115).

Modern chatbots such as ChatGPT, LaMDA and BERT are considered Large Language Models (hereinafter, LLM(s)) (Ge & Lai, 2023, p. 3). Basically, language models are statistical models that assign a probability to a sequence (e.g., a sentence) in order to compute the most likely follow up word (Carpenter & Altman, 2023, p. 2; Zhou et al., 2023, p. 1). Taking language models as mathematical representations of a language, all NLP models can be considered as language models (M. Zhang & Li, 2021, p. 831). ChatGPT was trained/programmed to create large deep neural network models (e.g., the Transformer architecture). By using the vast amount of information freely available on the internet, ChatGPT is capable of learning without supervision and generating human-like texts (Osmanovic-Thunström & Steingrimsson, 2023, p. 1; Zhou et al., 2023, p. 1). For instance, ChatGPT was trained based on the vast amounts of internet text data available on Wikipedia, Google, Reddit, Twitter, and other sources (Darlington, 2023, p. 50).

Before LLMs came into play, so-called traditional models (e.g., N-Gram) were quite common. These conventional language models use frequency-based probability estimations to compute the likelihood of a given sequence. However, these models suffer from low accuracy. To solve this problem, pre-trained LLMs such as GPT make use of the large information sources on the internet to build artificial neural networks. Through the usage of these networks, LLMs are capable to perform more complex mathematical calculations (such as derivable and differential methods), giving them the ability to compute more accurate probabilities and have a stronger context identification capability. Similar to traditional language models, LLMs compute the probability of a given sequence. However, LLMs produce vector representations of language segments as well, giving them an advantage compared to traditional models (M. Zhang & Li, 2021, pp. 831–832). While traditional language models mostly use supervised training, LLMs benefit from unsupervised training, and in the case of ChatGPT, from reinforcement learning from human feedback (hereinafter, RLHF) for finetuning the model (Dale, 2021, p. 114; Salvagno et al., 2023, p. 1; Taecharungroj, 2023, p. 2; M. Zhang & Li, 2021, p. 832).

Since supervised learning needs a vast amount of labelled data (which involves a lot of effort and expense), OpenAI introduced its first generative pre-trained language model (GPT-1) in 2018 (Radford et al., 2018, pp. 1–12; M. Zhang & Li, 2021, p. 832). Compared to traditional language models, this model was trained with huge amounts of unlabelled data and required minimal supervised training and finetuning on specific language-related tasks

(M. Zhang & Li, 2021, p. 831). After this initial success, OpenAI introduced GPT-2 in 2019 (Radford et al., 2019, pp. 1–24), which is "*a large unsupervised transformer language model with 1.5B parameters trained on 40GB of text"* (Dale, 2021, p. 114). This new model used a much larger amount of data and parameter scales compared to GPT-1 (M. Zhang & Li, 2021, p. 832). Given this large amount of text data, GPT-2 was able to perform various language-related tasks without the need of supervision, also known as zero-shot capability (Radford et al., 2019, p. 10), already indicating that the size of the model and the amount of data improves the performance and capabilities of LLMs (Zhou et al., 2023, p. 1). Although GPT-2 was already capable of generating comprehensive texts (with some exceptions), OpenAI decided not to publish the model for security concerns (Dale, 2021, p. 114).

Nevertheless, OpenAI continued to conduct research on its GPT series and announced GPT-3 in 2020. This third iteration of the GPT series was extended significantly by its model and data size, consisting of 175 billion parameters and 45 TB (vs. 40 GB in GPT-2) of text data (Dale, 2021, p. 115; M. Zhang & Li, 2021, p. 832; Zhou et al., 2023, p. 1). Because of this immense model scale-up, GPT-3 is capable of performing a wide variety of language-related tasks and can even accomplish newly tasks in zero-shot or in few-shot settings, using only a small sample data (Lund et al., 2023, p. 7; Zhou et al., 2023, p. 1).

Hence, OpenAI decided to launch ChatGPT, a generic or non-domain-specific LLM (Kung et al., 2023, p. 2), on November 30, 2022, which is based on a newly GPT-3 (also known as GPT 3.5), and made its browser-based intelligent chatbot freely available to the public (ChatGPT & Affairs, 2022, p. 379; Darlington, 2023, p. 49). Unlike common chatbots, ChatGPT is also referred to as conversational artificial intelligence interface, which was programmed by NLP models, ML (i.e., DL), and reinforcement learning. While ChatGPT is designed to have a conversation with users by understanding and generating natural language, the capabilities of the system go far beyond ordinary conversations. Thus, ChatGPT is not only able to respond to users' questions (so-called “prompts”) (Hacker, 2023, p. 2), but can also be used for creating new texts, such as stories, poems, and emails. Even summaries, translations and creating or correcting computer code are possible use cases (Arif et al., 2023, p. 1; Lund et al., 2023, p. 6; Salvagno et al., 2023, p. 1; Tlili et al., 2023, p. 2). In addition, by making use of DL and reinforcement techniques, ChatGPT is constantly improving due to the large number of users and their input i.e., feedback (Lund et al., 2023, p. 4; Salvagno et al., 2023, p. 1). Following these new technological possibilities, many researchers and professionals agree that ChatGPT could have a significant impact on the lives of many people, on almost every industrial sector, and on the labour market, which led to a lot of hype around ChatGPT (Lund et al., 2023, p. 5; Rana, 2023, p. 7; Taecharungroj, 2023, p. 2; Zhou et al., 2023, p. 4).

In fact, the hype (at the time of writing this study) of ChatGPT was so widespread and the development so rapid that OpenAI presented GPT-4 only a few months later in March 2023 (ChatGPT & Affairs, 2022, p. 380). This new GPT version is considered a multimodal model that accepts not only text as input, but also images (OpenAI, 2023a, p. 1). GPT-4 is available for ChatGPT Plus subscribers (paying customers) with a usage limit. In the future, OpenAI will try to offer GPT-4 for trial to users without a paid subscription, depending on capacity constraints (*GPT-4*, n.d.)[1]. However, the following sections (and this work) are based on the freely available GPT-3 i.e., GPT 3.5.

[1] Preprint

Capabilities

As it was mentioned above, ChatGPT is not an ordinary "question and answer" system. In fact, the conversation between user and system is based on past conversations as well as the current context. Furthermore, the responses provided depend on the accuracy of the instructions (i.e., prompts) given and are improved by RLHF. ChatGPT is not only capable of continuous self-improvement, but also admits mistakes and avoids answering inappropriate questions that pose a security risk to society and politics or contradict basic moral principles. ChatGPT also asks counter questions (when in doubt) to determine the exact intent of the user and reply to his/her question as accurately as possible. Eventually, the chat system also provides some insightful explanations for the answers given (Taecharungroj, 2023, p. 2; Zhou et al., 2023, p. 2).

Examining the potential use cases of ChatGPT reveals an abundance of seemingly boundless possibilities. A major use case of ChatGPT is seen in the generation of creative texts. Being trained on an enormous amount of textual web data, ChatGPT can be used for a variety of creativity-demanding, and language-related tasks. These include, for example, writing stories, poems, news articles, songs and many more. Moreover, ChatGPT can perform tasks that go beyond creative writing. For instance, the AI system can write computer code and provide explanations for it, it can create graphs, tables, and charts if correct instructions are given, and it can be used to explain complex concepts (Darlington, 2023, p. 50; Taecharungroj, 2023, p. 5; Zhou et al., 2023, p. 2).

All these application areas make ChatGPT a powerful tool for a wide variety of scenarios. Hence, OpenAI, with Microsoft as a big player in the tech sector on its side, has already integrated ChatGPT into the Bing search engine, potentially marking a new era for search engines. With this integration, users are now able to receive instant replies to their questions and are even provided with citations and the web page links from which the response information originates. Google, the market leader in search engines for years (*Global Search Engine Desktop Market Share 2023*, n.d.), also released its own chatbot "Bard" only a few weeks later, which can also be integrated into search engines (Zhou et al., 2023, p. 5).

Classification

Based on the design dimensions and the information provided about ChatGPT, a classification of the chatbot was carried out.[1]

In terms of **Prescriptiveness**, ChatGPT is best categorized as a non-task-oriented and generative-based chatbot. Unlike task-oriented chatbots that have specific functions and are designed to perform well on a range of specific tasks, ChatGPT is more focused on engaging in natural conversations with its users. There are no predefined actions or tasks to be performed, ChatGPT rather aims to find creative and context-appropriate responses.

Regarding the design dimension **Knowledge Base**, ChatGPT belongs to the category of generic chatbots. Unlike chatbots with domain-specific knowledge bases that are tailored to specific topics or industries, ChatGPT does not rely on a limited knowledge base. Instead, it has been trained on a huge amount of different (internet) data sources and is continually learning, so that it can respond to a wide range of user inputs and queries.

The dimension **Service** describes the type of relationship between the chatbot and the user. In the case of ChatGPT, it can be best categorised as an intrapersonal chatbot. This means that ChatGPT tailors its responses based on individual interactions and adapts its

[1] The classification according to the chatbot design dimensions was carried out by the author of this paper and is solely based on the aforementioned information on ChatGPT.

output to previous (or current) conversation(s) with a particular user. It does not provide a one-size-fits-all response, but adapts its responses based on the context it has learned from the ongoing conversation.

Referring to the **Response Generation** design dimension, ChatGPT is a machine learning-based model. Its ability to generate responses stems from extensive training on large datasets, allowing the chatbot to constantly adapt and improve its output. The model ChatGPT is using learns patterns and structures from the data it has been exposed to and applies this knowledge to generate coherent and contextual responses.

As far as the **Interaction** design dimension is concerned, ChatGPT is limited to text-based input only. Users interact with ChatGPT primarily through written text and ChatGPT responds in the same way. While there are advances in multimodal chatbots (e.g., GPT-4) that can handle different types of input such as images and speech, ChatGPT's primary mode of interaction remains text based.

As for the **Integration** dimension, ChatGPT is versatile in its accessibility. It can be accessed via a dedicated website or app and offers a straightforward user experience. In addition, ChatGPT can be integrated into other websites, apps and programmes i.e., plugins via APIs (Application Programming Interfaces) (*OpenAI Platform*, n.d.). This flexibility allows developers to use ChatGPT's capabilities in their own platforms and services to enhance the user experience and add conversational interfaces to various applications.

Finally, with regard to the dimension **Human Aid**, ChatGPT can be classified as an autonomous chatbot. It acts autonomously during the conversation process and relies mainly on its pre-trained data. While some control mechanisms are implemented to guide its responses and behaviour (and prevent malicious exploits), ChatGPT is not directly supervised by human intervention during the actual conversation with users.

In summary, ChatGPT has the characteristics of a non-task-oriented and generative-based chatbot with a generic knowledge base. It interacts with users on an intrapersonal level and adapts responses based on previous conversations and the current context. Its response generation is based on machine learning and is mainly text based. Furthermore, ChatGPT can be easily integrated into different platforms via APIs, offering developers a wide range of possibilities to enhance their own applications or services. Finally, ChatGPT operates autonomously and relies on its pre-trained data during the conversation. Understanding these design dimensions provides valuable insights into ChatGPT's capabilities and limitations as an AI chatbot.

It should be mentioned that, to the best of the author's knowledge, such a classification according to the chatbot design dimensions of ChatGPT has not been carried out in any scientific paper.

Limitations

Even though ChatGPT has improved considerably in many respects compared to its predecessors, the system still demonstrates some weaknesses. Thus, ChatGPT's pre-trained data at this time is based on sources up to 2021 and consists predominantly English data (Darlington, 2023, p. 51; Floridi & Chiriatti, 2020, p. 685; Rana, 2023, p. 8), which from time to time may lead to misinformation (Zhou et al., 2023, p. 4). Furthermore, LLMs suffer from so-called "hallucinations". This term refers to responses from the system that appear specific and correct, but are in fact completely wrong (Ge & Lai, 2023, p. 5; Zimmerman, 2023, p. 3). Teubner et al. demonstrate, for example, that when searching for scientific articles with the help of ChatGPT, mostly wrong titles, citations and DOI are returned (Teubner et al., 2023, p.

97). Another disadvantage of algorithmic systems trained by internet data is referred to as "algorithmic bias". This type of bias describes discriminating output, which can be related to gender, race, religion, or political orientation, for example. Not only the data used to train ChatGPT, but also conscious or unconscious decisions made by the programmers in the design process are possible reasons for such a bias (Hacker, 2023, p. 2; Rozado, 2023, p. 1). For instance, Rozado shows in an experiment with the responses of ChatGPT and various political orientation tests he conducted that the chatbot is predominantly politically left-leaning (Rozado, 2023, p. 4). In addition, in a recent study, it was shown that ChatGPT provides different answers to different users despite the same input given, which can lead to unfair access to correct information (Tlili et al., 2023, p. 20). Another drawback of ChatGPT is its robustness. While ChatGPT has been trained to present harmless responses, there are ways around these "boundaries". So-called “instruction attacks” or “prompt injections”, for example, cause the system to behave in ways it was not designed to behave, resulting in unethical or even dangerous responses. In addition, since ChatGPT answers are based on statistical methods, definite answers (e.g. in mathematics) can be answered incorrectly (Zhou et al., 2023, p. 4). Interestingly, only one study looking at the use of ChatGPT for management scientists and its impact claims that the chatbot lacks creativity (Rana, 2023, p. 7), which is exactly the opposite of the opinions of other studies (Darlington, 2023, p. 50; Taecharungroj, 2023, p. 5; Zhou et al., 2023, p. 2).

Possible Implications

Even if, as just described, ChatGPT still presents some flaws, researchers generally agree that the new chatbot (and other similar technologies) will have far-reaching effects on our society and everyday life. The following paragraphs describe some of these possible consequences.

For instance, research and the public believe that ChatGPT could have a profound impact on the educational sector. As a powerful writing tool, there is a legitimate chance that students in the future will use ChatGPT to mindlessly get their written work done and lose their ability to come up with new ideas (Arif et al., 2023, p. 1; Darlington, 2023, p. 51). Others are more positive about the development and believe that ChatGPT could act as a personal virtual tutor and help students understand new or complex concepts (Darlington, 2023, p. 50). Recently, ChatGPT was even reported to be able to almost pass the USMLE (United States Medical Licensing Examination) without prior training (Kung et al., 2023, p. 8). These potential developments, and the remarkable result of ChatGPT at the USMLE, as an example of the new chatbot's capabilities, lead to the assertion by some researchers that the education sector (teachers and students alike) needs to build more skills and adopt new student assessment methods (Taecharungroj, 2023, p. 8; Tlili et al., 2023, p. 22).

Another "hot topic" in the literature is the use of ChatGPT in science and its related concerns regarding academic misconduct and plagiarism. Moreover, scientists currently discuss about the accountability of the chatbot's content in academic literature and the question if ChatGPT should be listed as a (co-) author (Arif et al., 2023, p. 1; Floridi, 2023, p. 4; Osmanovic-Thunström & Steingrimsson, 2023; Zhou et al., 2023, pp. 4–5). Regardless of this discussion, some researchers have already used ChatGPT for their scientific publications, either listing the chatbot as an author, e.g., (ChatGPT & Affairs, 2022), or by mentioning ChatGPT in the acknowledgements/disclaimer, e.g., (Floridi, 2023, p. 7). Despite these ethical concerns, Lund et al. assume that ChatGPT could increase the productivity and help researchers in the citation process (Lund et al., 2023, p. 13).

However, productivity boosts do not only affect the scientific sector and could also have negative consequences. Taecharungroj for example used the latent dirichlet allocation (LDA) topic modelling algorithm to analyse the commentaries of ChatGPT early adopters on Twitter. The author suggests that many users are optimistic about the transformative impact of ChatGPT on their daily work, especially in the marketing field. Many believe that tasks in product design, content creation and copywriting are made significantly easier through the use of ChatGPT. Nonetheless, many first-time users also express their concern that the new chatbot could lead to massive job cuts, especially regarding white-collar and creative professions (Taecharungroj, 2023, pp. 7–8). Zhou et. al share this opinion: ChatGPT as a "universal assistant" could have a major impact on pretty much every industry, *"including education, mobile, search engine, content production, and medicine"* (Zhou et al., 2023, p. 4), resulting in jobs that will disappear and new job profiles being created (Floridi, 2023, p. 5). Another concern with automated text generation is the fast, targeted and cheap creation and spread of fake news, which could even endanger the world's democracies (Hacker, 2023, pp. 2–3; Renn, 2023, p. 1).

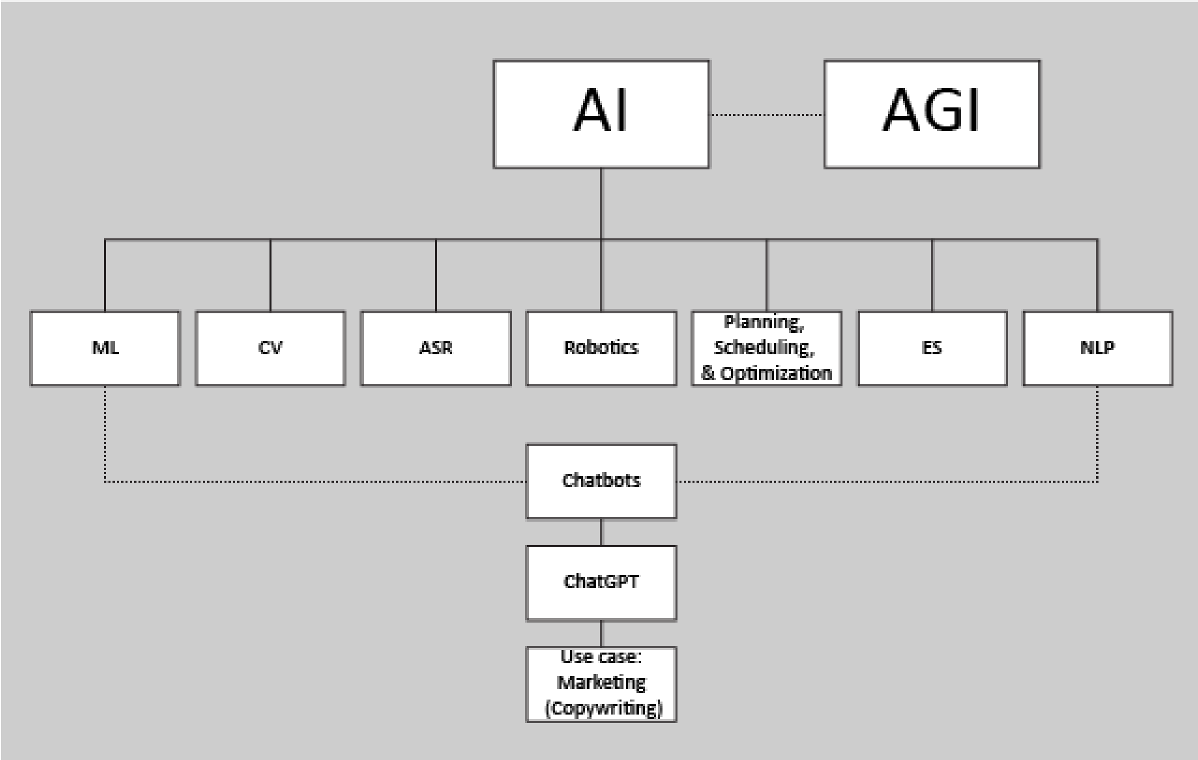
Due to these threats posed by the rapid development of AI and chatbots, researchers are proposing ethical and legal frameworks, an AI governance, for the safe use of artificial intelligence (Taecharungroj, 2023, p. 8; Zhou et al., 2023, pp. 4–5). While AI development is likely to accelerate due to multi-million investments and trend applications such as ChatGPT, the European Union is still working on a legal framework, the so-called EU AI Act. The proposed EU AI Act is a set of legislation aimed at regulating the use of AI in the European Union. One of the AI Act's core clauses concerns General-Purpose AI Systems (GPAIS), which are described as models capable of performing well on a wide range of activities for which they were not expressly trained. This definition would therefore also apply to ChatGPT (Hacker, 2023, p. 1).

Given the topicality, scholarly work related to ChatGPT is limited and there is often no scientific evidence yet to support the opinions expressed by researchers and the public, e.g., in the field of marketing. Although chatbots do not represent an innovation, recent academic papers have tended to place more emphasis on the technical aspects of chatbots (Taecharungroj, 2023, pp. 1–3).

At the time of writing this work, to the best of the author's knowledge, no academic work has been devoted to the use of ChatGPT in marketing, or more precisely, in copywriting.

The graphic below provides an overview of the structure of this work.

Figure 1 - Structure of this work



Note. Figure adapted from (de-Lima-Santos & Ceron, 2021, p. 19). "AGI" and "Chatbots" are connected with dashed lines because AGI is still under development and chatbots can use other AI fields than ML (i.e., DL) and NLP. However, this work focuses on the chatbot ChatGPT, which is mainly based on ML and NLP.

Marketing AI

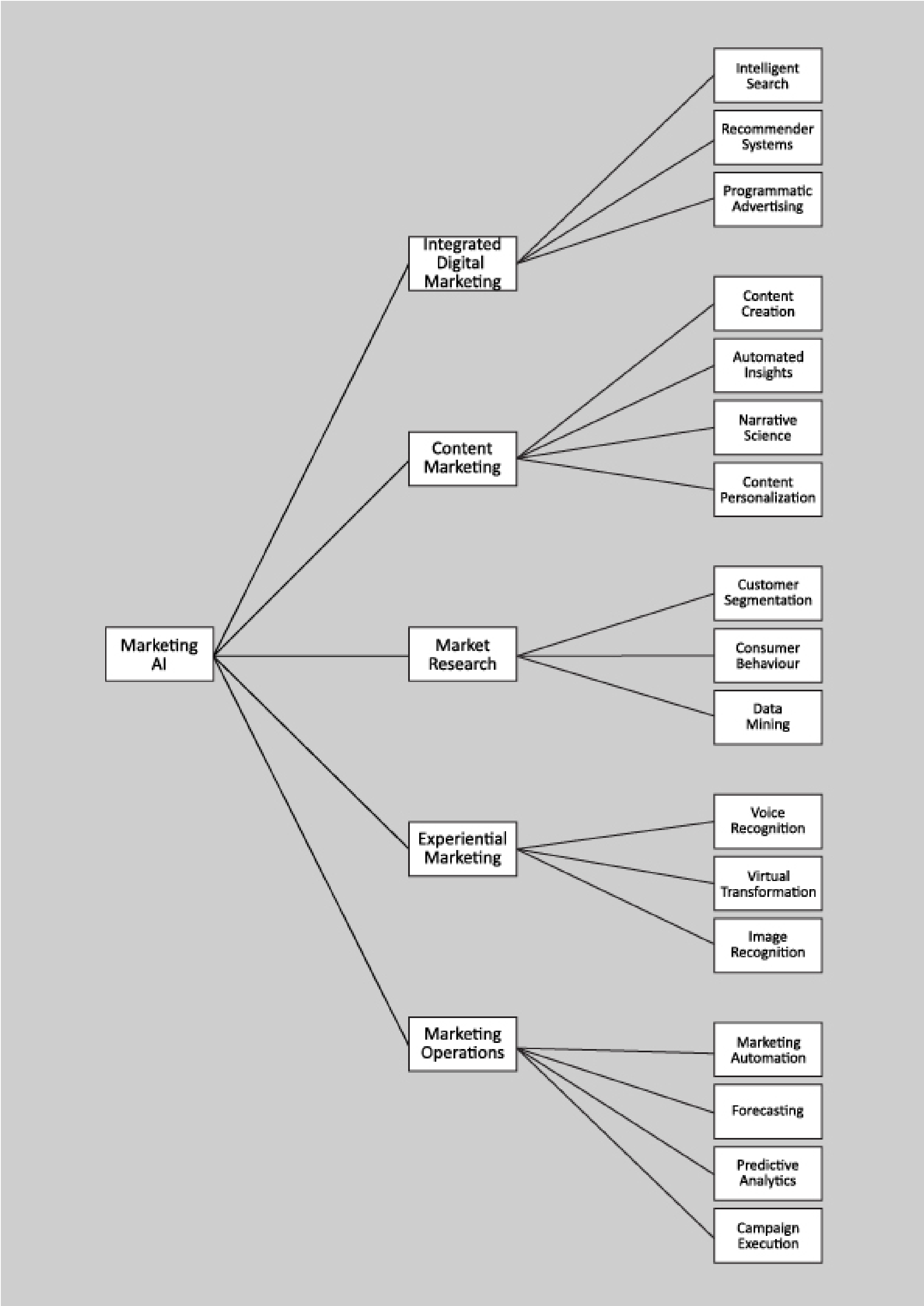
Although academic literature on the use of ChatGPT in marketing is virtually non-existent due to its recent nature, AI has been an important tool for marketers for some time now. In marketing, AI leverages customer data to make predictive statements about customer behaviour and improve the customer experience (Sumitha, 2022, p. 16). Overgoor et al. define marketing AI as *"the development of artificial agents that, given the information they have about consumers, competitors, and the focal company, suggest and/or take marketing actions to achieve the best marketing outcome"* (Overgoor et al., 2019, p. 157). Compared to traditional marketing, marketing AI is based on the aggregation and analysis of vast amounts of data that facilitate "one-to-one marketing", also known as micro segmentation (Sumitha, 2022, pp. 17–18).

It was already successfully predicted in 2015 that 20 per cent of all business content will be written by machines in 2018 (Gartner, 2015). Moreover, a longitudinal study conducted by Bünte in Germany, Austria, and Switzerland shows that 93 percent of marketing managers surveyed see AI as an important technological tool for marketing purposes (Bünte, 2021, p. 32). These exemplary figures illustrate how important AI has become for marketing and businesses in general and, for several reasons, it can be assumed that the use of marketing AI will increase in the coming years.

One major reason why the use of AI in marketing has risen sharply and is likely to continue to rise are efficiency improvements. This factor is mentioned throughout the research and literature: Users and companies are likely to enjoy efficiency improvements with the use of AI, e.g., (Qin & Jiang, 2019, p. 338; Sumitha, 2022, p. 20; Volkmar et al., 2021, p. 367). Because AI is capable of processing and analysing large amounts of data, it has the power to support companies in important decision-making processes and automize many tasks. However, not only the increase in efficiency is a decisive factor for the success of marketing AI, but the increase in effectiveness plays a central role as well. In fact, many insights into the data collected by companies and the decisions based on it have only become possible through the use of AI (Sumitha, 2022, p. 20; Volkmar et al., 2021, p. 368). Using AI in marketing, it is possible to identify correlations between data points and gain valuable insights into the customer base. In addition, these data insights allow tailored advertising, which compared to traditional advertising, is more cost-efficient (Sumitha, 2022, pp. 17–18). Finally, the collection and analysis of customer-related data allows companies i.e., marketers to anticipate customer behaviour and improve innovation capabilities in marketing strategies and advertising campaigns (Chintalapati & Pandey, 2022, p. 57; Volkmar et al., 2021, p. 368).

However, it should be mentioned that AI is not yet in use in all marketing fields and the state of research on the respective fields also varies. Before going into more detail on this topic, the following graphic illustrates the themes & sub-themes of marketing AI.

Figure 2 - Marketing AI Themes and Sub-Themes



Note. Figure adapted from (Chintalapati & Pandey, 2022, p. 52)

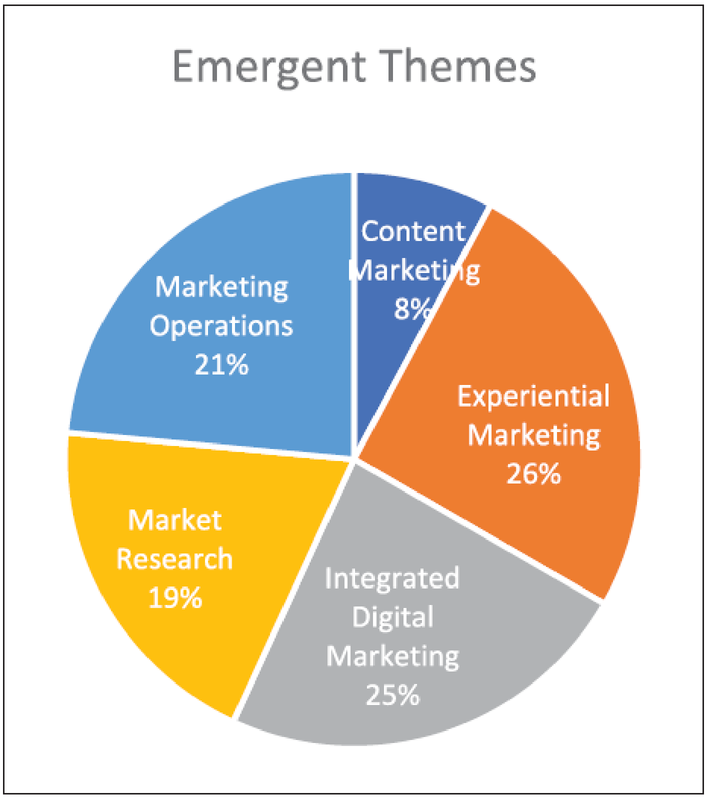
In 2022, Chintalapati and Pandey conducted a systematic literature review on the use of AI in marketing. In their analysis, the authors identified five distinct marketing AI themes that are currently being explored in academia. These five superordinate themes of marketing AI are "Integrated Digital Marketing", "Content Marketing", "Market Research", "Experiential Marketing", and "Marketing Operations". Furthermore, the study reveals 17 sub-themes (so-called activity levers) and 170 different use cases of AI applications in marketing (Chintalapati & Pandey, 2022, p. 43).

Comparing the body of published literature, it appears there is a discrepancy between the theory and practice of marketing AI. Today, marketing professionals make substantial use of AI for operational processes, such as the segmentation of customers. More specifically, the leading applications of marketing AI can be found in content creation (Theme: "Content Marketing"), voice search (Theme: "Experiential Marketing"), as well as predictiveness analysis, lead scoring, ad targeting, and dynamic pricing (Theme: "Marketing Operations")[1]. Nevertheless, AI is still relatively rarely employed to make strategic marketing decisions (Sumitha, 2022, p. 18; Volkmar et al., 2021, p. 367). When compared with the state of research in the various marketing AI themes, similarities but also differences become apparent. While the themes "Experiential Marketing", "Integrated Digital Marketing", and "Marketing Operations" are undergoing a considerable amount of research efforts, "Content Marketing" is still little explored (Chintalapati & Pandey, 2022, pp. 54–55). The chart below provides an overview of the scope of marketing AI research in the various themes.

---

[1] The allocation of the AI applications to the five superordinate marketing AI themes was conducted by the author of this paper.

Figure 3 - Marketing AI and its research coverage in different marketing AI themes



Note. Figure adopted from (Chintalapati & Pandey, 2022, p. 56)

The chart above illustrates the limited scientific coverage of the marketing AI theme "Content Marketing". Therefore, as a sub-theme of content marketing, the following subsection is dedicated to copywriting.

AI & Copywriting

As briefly mentioned before, some researchers and early adopters have made the statement that ChatGPT can be used for creative-demanding and language-related tasks, such as content production, and more specifically, for copywriting (Taecharungroj, 2023, pp. 7–8; Zhou et al., 2023, p. 4). Other authors share the same opinion and claim that AI and applications like ChatGPT are beneficial for copywriting. Thus, Deng et al. state that copywriters are freed from the monotony of repetitive work. Since professionals have to write very similar copies for different settings i.e., formats such as newsletters, social media posts, and websites, AI can be a solution to avoid this repetitive work and save time (Deng et al., 2019, p. 363). Furthermore, in the age of (infinite) information availability, copywriters can easily get overwhelmed by the information they have to process (so-called "information overload"). In addition, some products may require specialist training and tutoring for copywriters, to be able to write about them (Deng et al., 2019, p. 357; X. Zhang et al., 2022, p. 12423). Therefore, a self-learning AI that can access a huge database and generate content in seconds is likely to be beneficial.

This paper classifies copywriting as part of the superordinate theme content marketing, see also (Farnworth, 2015). In general, the term content encompasses both static content and dynamic rich media content such as videos, podcasts, newsletters, blogs, and other content formats on websites (Holliman & Rowley, 2014, p. 2; Köse & Sert, 2016, p. 838; Van Noort et al., 2020, p. 414). Content marketing refers to *"a strategic marketing approach focused on creating and distributing valuable, relevant, and consistent content to attract and retain a clearly defined audience — and, ultimately, to drive profitable customer action"* ('What Is Content Marketing?', n.d.).

Copywriting, as a form of advertising content, can be further categorized into communication copy and sales copy. The main purpose of communication copy is to enhance brand awareness through advertising, with a focus on growing brand influence, building brand image, and promoting a corporate culture. Thus, communication copy is not intended to directly promote sales through advertisements. Sales copy, on the other hand, is explicitly concerned with promoting sales through advertising, with the goal of increasing revenue from direct sales (Deng et al., 2019, p. 361).

In the scientific field, Arthur Kover in particular has been involved with the subject of copywriting. In 1995 and 1996, he investigated if copywriters have implicit theories to explain their work. Based on in-depth interviews with copywriters, he suggests an implicit theory, which consists of two steps: **(1)** Breaking through, and **(2)** Message delivery. Breaking through refers to the customer's interest that should be sparked by the advertising efforts. According to Kover, there are two methods to break through i.e., capture the attention of the customer: Subversion (surprising someone with something unexpected or unusual) and forcing (e.g., using sheer force of words). Once the attention is grabbed, there is a small window of opportunity for the advertiser to get the message delivered (Kover, 1995, pp. 599–601, 1996, pp. 9–10). One way of capturing the attention of consumers is seen in advertising creativity.

Advertising Creativity

In fact, advertising creativity mainly aims to draw attention and meet specific objectives set by external others, which distinguishes it from creativity in the arts. Measuring the success of advertising creativity and creativity in the arts also differ. In the arts, success is typically measured by the (subjective) aesthetic appeal derived from the creative work. On the contrary, advertising creativity is only considered successful if it meets two goals: The first goal is to grab the consumers' attention, and the second is to trigger a certain effect i.e., behaviour (e.g., purchase intention) in them (El-Murad & West, 2004, p. 190). As becomes quite clear here, the two goals of advertising creativity are quite similar to the two components of the implicit theory of copywriters described by Kover (Breaking through and Message delivery). Therefore, the following paragraphs will look more closely at theories of advertising creativity.

In general, research suggests that creativity is an important factor in advertising as it can affect persuasion and thus consumer behaviour (Campbell et al., 2022, p. 26; El-Murad & West, 2004, p. 188; Kover, 2016, p. 236; Smith et al., 2007, p. 819; Smith & Yang, 2004, p. 31; West et al., 2008, p. 35). In fact, *"creative advertising may actually bestow value to the advertised brand"* as it facilitates recall on an unaided basis (Till & Baack, 2005, p. 55).

Kover describes creativity as a type of innovation that is subject to certain boundaries. However, unlike other innovations that have the potential to render previous accomplishments useless, advertising creativity seldom leads to destruction. Moreover, the judgement about whether something is creative depends on the viewer. For instance, the evaluation of a work differs when asking creative professionals, the brand's clients, or the general public about the creativity at hand (Kover, 2016, p. 235; Smith & Yang, 2004, p. 32). Consequently, consumers interpret (creative) advertising based on their needs and preferences, and their evaluation may differ greatly from that of creative minds (West et al., 2008, p. 35).

Although advertising creativity plays an important role in the advertising industry, there seems to be a lack of a universally accepted definition of the term in the literature. However, the authors Smith et al. claim that most definitions are similar to the one provided by Leo Burnett (Smith et al., 2007, p. 819). Burnett defines advertising creativity as *"the art of establishing new and meaningful relationships between previously unrelated things in a manner that is relevant, believable, and in good taste, but which somehow presents the product in a fresh new light"* (El-Murad & West, 2004, p. 190). (Advertising) Professionals, on the other hand, often answer the question "What is creativity?" with something similar to "I know it when I see it" (Kover, 2016, p. 235).

Nevertheless, research has agreed on two key aspects i.e., components of advertising creativity: **(1)** Divergence, and **(2)** Relevance. Divergence is the main factor that contributes to advertising creativity and describes advertising elements that are supposed to be novel i.e., original, different, or unusual in some way. Although divergence mainly contributes to the perceived creativity of advertising, relevance also plays a central role as it interacts with divergence (Smith et al., 2007, p. 829). Hence, an advertisement that only contains divergent elements is insufficient. To be judged creative by the target group, the advertising must also be relevant (or: appropriate, meaningful) and capable of addressing a specific problem or attaining a desired goal (Campbell et al., 2022, p. 27; El-Murad & West, 2004, p. 189; Smith et al., 2007, p. 820; Smith & Yang, 2004, pp. 34–36).

Both elements, divergence and relevance, as well as advertising creativity itself, are characteristics of an advertisement that are difficult to measure and, as already mentioned, depend on the subjective evaluation of the viewer. Therefore, the authors Smith et al. have dedicated a study to modelling and measuring advertising creativity. The authors' findings suggest that when measuring the two components of advertising creativity, it is crucial to use verified scales i.e., factors that are the determinants of each component. After conducting several pre-tests and a large-scale study, they identified the following factors of divergence and relevance. To measure (ad) divergence, the five factors "originality", "flexibility", "synthesis", "elaboration", and "artistic value" (more on these factors in **Table 2**) should be considered. And to measure (ad) relevance, the factors "ad-to-consumer relevance" and "brand-to-consumer relevance" should be included (Smith et al., 2007, p. 829).

"Ad-to-consumer relevance" is accomplished by advertisement execution elements (text, music, images, etc.) that establish a meaningful link to consumers. As an example, the authors mention the use of Beatles music in advertising, which can establish a meaningful connection with Baby Boomers. “Brand-to-consumer relevance”, on the other hand, occurs when there is a meaningful connection between the brand (or product category) and the consumer (e.g., by showcasing the brand in situations familiar with the consumer’s everyday life). A third factor, “ad-to-brand relevance”, describes the degree of alignment between the ad and the brand (Smith et al., 2007, pp. 820–821). However, as “ad-to-brand relevance” is not a significant predictor of advertising creativity, the authors did not include it in their final model (Smith et al., 2007, p. 829).

Finally, it should also be mentioned that some researchers assume that not all advertising needs to be creative. When it comes to stable, well-established products, so-called “low-involvement products” like toilet paper, mere reminders may be adequate (Vakratsas & Ambler, 1999, p. 33).

As described above, creativity is an important advertising element that copywriters can use to potentially influence the perception and thus the behaviour of consumers.

While some researchers claim that AI and applications like ChatGPT will make the job of the professional copywriter (and other creative jobs) obsolete, e.g., (Coffin, 2022, p. 614; Floridi, 2023, p. 5), there is no scientific evidence to support these statements. Hence, the question of whether ChatGPT is only a helpful tool for copywriters or will replace them altogether remains open. In addition, the fact that content marketing is the least researched of all five functional themes of marketing AI was highlighted, which underscores the importance of this study. While the use of AI for automatic text generation in the news sector is abundantly researched, see for example (Broussard et al., 2019; de-Lima-Santos & Ceron, 2021; Kotenidis & Veglis, 2021; Melin et al., 2018; Noain-Sánchez, 2022), research on the use of AI for copywriting purposes is scarce.

Based on the two subsections just presented, the question arises whether consumers perceive texts written by ChatGPT as creative as those of their human counterparts. As described in the next chapter, attitudes towards AI could also have an impact on the evaluation of AI-generated content. This leads to the following research question:

How do consumers perceive the creativity of copy generated by ChatGPT compared to copy by human copywriters, and how do their attitudes towards AI in terms of creativity affect this perception?

To the best of the author's knowledge, there is no scientific paper dedicated to this question. Although this question cannot be directly equated with the effectiveness of advertising, it could provide important indications of consumers' perceptions of AI-generated advertising copy. With the EU AI Act still in process and researchers, such as Hacker, requiring disclosure of users who use LGAIMs (Large Generative AI Models) for content creation (Hacker, 2023, p. 6), the scientific and managerial relevance of this research question is given.

Chapter 2: Theoretical Background & Hypotheses

In the past, scholars have already addressed the perception of AI-generated output (not specific to copywriting) and of machines in general. According to Sundar, people view the decisions made by algorithms as efficient and objective, making them well-suited for mechanical tasks but ill-suited for tasks involving subjective judgements and emotional abilities (Sundar, 2020, pp. 79–80). Thus, some abilities or characteristics of AI are desired (referred to in the literature as “algorithmic appreciation”), and others are overwhelmingly rejected by humans (referred to in the literature as “algorithm aversion”) (Agudo et al., 2022, p. 2). In another study devoted to so-called "algorithmic journalism" Kotenidis and Veglis examined the extent to which the use of AI has changed or will change the job profile of journalists. Their study also discusses whether AI could soon replace the job of the (human) journalist. The authors conclude their article that AI serves as a complement to, not a replacement for, journalists because it lacks creative skills (Kotenidis & Veglis, 2021, p. 252). As mentioned earlier, Rana made the same statement regarding ChatGPT (Rana, 2023, p. 7). In general, some scholars argue that machines cannot be as creative as humans are as they lack the ability to create “truly original” content (Van Noort et al., 2020, p. 415). Hypothesis no. 1 is therefore:

H1: Participants' evaluations will indicate a higher level of perceived creativity for human-written copy compared to AI-generated copy, suggesting a positive impact of human intervention on creative output.

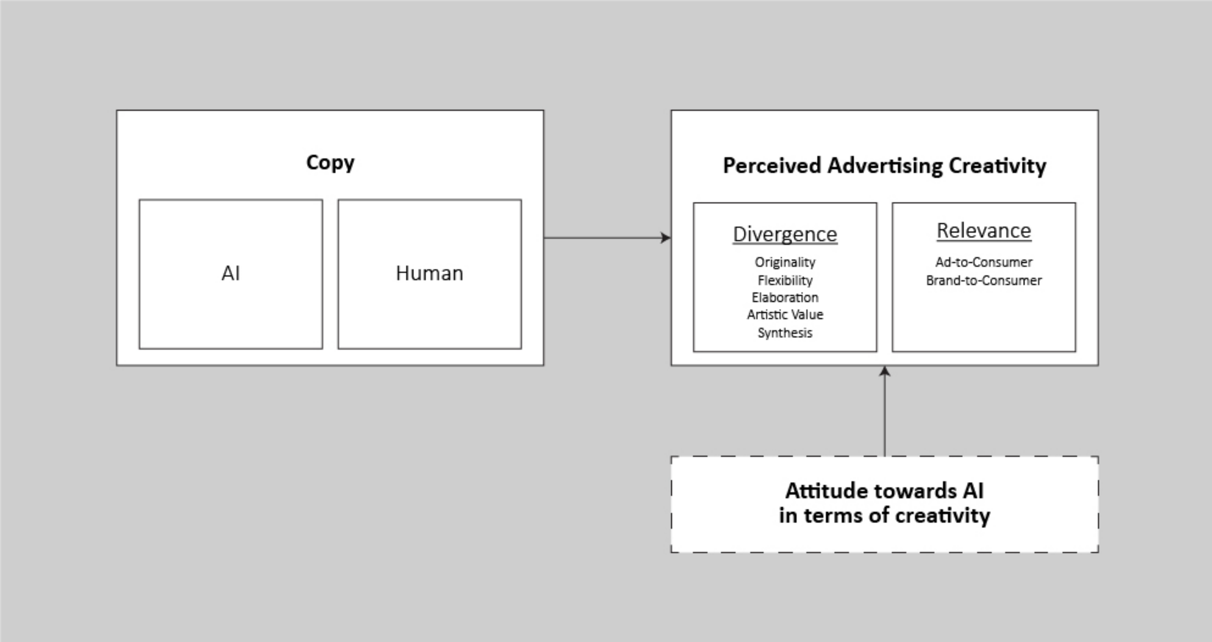
Another recent study by Hong et al. examined, among other research objectives, the relationship between the belief about creative AI (i.e., hold the view that AI can be creative) and AI-generated music. They found a positive relationship between people's perception and their belief about AI's creative ability, meaning *"a predisposition to be open to AI predicts enjoyment of its products".* In other words, results suggest that the assessment of creative work is biased by perceptions of AI, rather than solely relying on the output's quality (Hong et al., 2021, p. 3). Accordingly, if people consider creativity as a human-only characteristic, they may depreciate AI-generated copywriting because of its source. Hypothesis no. 2 is therefore:

H2: There is a positive relationship between the belief that AI can be creative and the evaluation of its texts.

Note. Adapted from (Hong et al., 2021, p. 4)

The research objective is thus to determine, on the one hand, consumer perceptions of AI-generated copy (vs. human-generated copy) in terms of creativity, and on the other hand, the influence of attitude towards (creative) AI on consumer perceptions of AI-generated copy. Based on the assumptions (i.e., hypotheses), the research objective, and the underlying theory (see **Advertising Creativity**), the following conceptual framework was developed.

Figure 4 – Research Hypothesis Model



Note. Figure adapted from (Ananthakrishnan & Arunachalam, 2022, p. 6296; Smith et al., 2007, p. 823)

As can be seen in the note below the figure, the theoretical model of Ananthakrishnan and Arunachalam was partly adapted. Similar to this study, the authors investigated consumer perceptions between AI- and human-generated content in terms of creativity. Nevertheless, there are some apparent differences to this study. First of all, it should be mentioned that the authors have focused on brand content. Unlike product advertising or sales copy, the aim of brand content is to convey the values of a brand (Ananthakrishnan & Arunachalam, 2022, p. 6294). A second distinctive factor is the way in which the authors measured creativity. Thus, for the construct "content creativity", they rely on the measurement of variables adopted from an earlier study by Smith and Yang, see (Smith & Yang, 2004), and assumed that *"there is only little conceptual development in terms of divergence in advertising literature (only one construct: originality is considered)"*

(Ananthakrishnan & Arunachalam, 2022, p. 6295). However, as mentioned in the previous chapter (see **Advertising Creativity**), Smith and Yang (and colleagues) developed a validated model including scales for measuring divergence and relevance, and hence, overall advertising creativity, in 2007. Accordingly, the independent variable “Advertising Creativity” consists of both second-order composite latent factors “Divergence” and “Relevance”, which are jointly measured by five resp. two first-order factors i.e., determinants (Smith et al., 2007, p. 823). One remaining difference in this study compared to the study by Ananthakrishnan and Arunachalam, is the examination of the consumer attitude towards AI (and its impact on consumer perception of AI-generated copy) in this study. For these reasons, it is considered that this study differs significantly from that of Ananthakrishnan and Arunachalam.

The following table provides a more detailed description of the individual components of the theoretical model.

Table 2 - Components Theoretical Model

Component	Type	Description	Citation
Perceived Advertising Creativity	Dependent Variable (IV)	See **Advertising Creativity**	-
• Divergence	Second-order composite latent factor	See **Advertising Creativity**	(Smith et al., 2007, p. 823)
○ Originality	First-order factor Divergence	*"Ads that contain elements that are rare, surprising, or move away from the obvious and commonplace"*	(Smith et al., 2007, p. 821)
○ Flexibility	First-order factor Divergence	*"Ads that contain different ideas or switch from one perspective to another"*	(Smith et al., 2007, p. 821)
○ Synthesis	First-order factor Divergence	*"Ads that combine, connect, or blend normally unrelated objects or ideas"*	(Smith et al., 2007, p. 821)
○ Elaboration	First-order factor Divergence	*"Ads that contain unexpected details, or finish and extend basic ideas so they become more intricate, complicated, or sophisticated"*	(Smith et al., 2007, p. 821)
○ Artistic Value	First-order factor Divergence	*"Ads that contain artistic verbal impressions or attractive colors or shapes"*	(Smith et al., 2007, p. 821)
• Relevance	Second-order composite latent factor	See **Advertising Creativity**	(Smith et al., 2007, p. 823)
○ Ad-to-Consumer Relevance	First-order factor Relevance	See **Advertising Creativity**	(Smith et al., 2007, p. 823)
○ Brand-to-Consumer Relevance	First-order factor Relevance	See **Advertising Creativity**	(Smith et al., 2007, p. 823)
Human/AI-generated Copy	Independent Variable (IV)	Copy written by either Human or ChatGPT	-
Attitude towards AI in terms of creativity	Predictor Variable	Possible factor influencing the evaluation of AI-generated copy	-

Note. This table was created by the author of this paper.

Chapter 3: Methodology

In order to obtain an answer to the research question and to be able to test the hypotheses, an online questionnaire survey will be conducted. This method was chosen as a high standardization and a low interviewer bias (influence by interviewer on participants) is provided (Albers et al., 2009, pp. 51–52). In addition, questionnaire surveys are particularly well-suited for studies revolving around individual persons (here: consumers) and for measuring people's preferences, beliefs and attitudes in an unobtrusive way (Bhattacherjee, 2012, p. 73). Furthermore, as two existing models were integrated into a new model i.e., an existing model was extended, a positivist research design, which aims to test theories (or hypotheses), is deemed appropriate (Bhattacherjee, 2012, p. 41). Other positivist research designs such as experiments (and quasi-experiments) were rejected because the independent variable (Human/AI-generated Copy) may not be easily manipulated or controlled (Bhattacherjee, 2012, p. 83).

To measure participants' ratings, established verbal 5-point Likert scales were used (Albers et al., 2009, p. 69). An odd number of values was chosen to allow participants to have a neutral opinion ("undecided") (Bhattacherjee, 2012, p. 48):

completely right	quite right	undecided	quite wrong	completely wrong
(5)	(4)	(3)	(2)	(1)

Due to time and cost restrictions, a full census was not an option for the data collection (Albers et al., 2009, p. 79). Therefore, a convenience sample was used. This method belongs to the non-probability procedures of sample determination and the author is aware that it does not lead to representativeness and/or generalizability (Albers et al., 2009, p. 83; Sarstedt et al., 2018, p. 652). Since it is not within the realm of possibility to identify i.e., determine all the potential samples that can be drawn from the population of inference, probability methods were rejected (Sarstedt et al., 2018, p. 651). The difference between theory application and effects application research should also be emphasised. While effects application research explores the effects of a specific theory in real-world scenarios, taking into account all possible external factors, including the sample structure, the intention of theory application studies is not generalizability, but rather the examination of specific effects in a defined research context (Sarstedt et al., 2018, p. 653). As this is a highly controlled study in which only copy texts (e.g., no visuals) are assessed and this is rarely the case in real life scenarios (exception: email subject lines, which may also contain emojis), this work is supposed to be seen as a theory application study and therefore a convenience sample is considered appropriate. As this is the first study of its kind and a convenience sample is used, it is considered a pilot test (Bhattacherjee, 2012, p. 69)

The participants were predominantly students and/or of younger age. They were provided with a link to the survey on Unipark. The online survey was conducted in English between the beginning and the middle of July 2023. In order to generate the advertising texts to be examined for this study, ChatGPT (GPT 3.5) was used on the one hand and one copywriter was hired on Fiverr, a freelance service provider, on the other. Copywriters on Fiverr were sorted by "Best Selling" and one of them with high ratings was selected.

Similar to a study by Agudo et al., participants were not told which AI tool specifically was used, as some participants might not be familiar with ChatGPT. Instead, Artificial Intelligence (AI) was simply named as one of the text sources (Agudo et al., 2022, p. 3). Since the interest of this research lies in the subjective assessment of a non-expert, no definitions were provided for the terms “Artificial Intelligence” and “Creativity” (Agudo et al., 2022, p. 4; Sai et al., 2022, p. 6).

It is known that in addition to human evaluation of automatically generated text, there are also untrained, and machine-learned metrics to assess the output of NLG. While untrained evaluation metrics assess the similarities between machine-generated texts and human-written texts using measures such as content overlap, string distance or lexical diversity, machine-learned metrics are trying to simulate the judgement of a human (Dong et al., 2023, pp. 29–30). However, as stated by Sai et al., current NLG metrics display weak correlations with human judgements, lack interpretability, and fail to capture detailed aspects of the given task (Sai et al., 2022, p. 4). Since consumers are the ultimate judges of advertising effectiveness (Linwan Wu & Jing Wen, 2021, p. 134), a human evaluation setup was deemed appropriate.

As can be seen in the table of contents of the book “THE COPYWRITING SOURCEBOOK” (Maslen, 2010), copywriting is ubiquitous on a wide variety of platforms (digital and print) and in different formats, e.g. in product brochures, emails, articles, on websites and more generally, in headlines. All these platforms and formats are intended to have an impact on consumer behaviour through targeted choice of words (Achar et al., 2016, p. 5). In addition to text, other formats, such as images, videos, and audio also play a central role in advertising messages and can evoke the desired consumer perception and behaviour (Januszewicz et al., 2022, pp. 3–4). Since it is likely that the use of other content formats, besides text, will also have an influence on the evaluation in terms of creativity, only human/AI text is used in the survey. More specifically, the (human) copywriter and ChatGPT were tasked with writing headlines promoting ice cream. Just as in the study by Kemp et al., who studied the emotional response to affect-laden advertising for hedonic products, a food item (ice cream) was selected due to food's general attractiveness (Kemp et al., 2012, p. 344).

All copywriters, ChatGPT and the Fiverr freelancer, received the following instructions:

Your task is to write three headlines for promoting ice cream. The copies should be creative, attention-grabbing, and persuasive, aimed at enticing potential customers to purchese the ice cream. There are no strict length requirements for the copies i.e., headlines.

Human copywriters received the following additional instructions:

One of your headlines (i.e., copies) will be used for research purposes only and treated with utmost confidentiality. By participating, you agree to the use of your copy for research and evaluation purposes. Thank you for your contribution and commitment to producing creative and effective advertising i.e., sales copy. Your input will contribute significantly to our understanding of consumer perceptions in advertising. Make sure that your copy is original, not plagiarised or generated by AI or similar technologies. Plagiarism is not permitted and can undermine the integrity of the assessment process.

In a randomized selection process, one text per source was chosen among a total of six proposed texts for further examination. Two exclusionary questions were added at the beginning of the survey to ensure participants like the food item ("*Do you like ice cream?*") and actually understand the survey (*"Do you know AI?"*). After assessing the attitude towards creative AI, each participant was assigned one AI-generated and one human-written ad copy i.e., headline. As the order in which the ad copy was presented could have an influence on the participants' evaluation, 50 percent of the participants were randomly presented with ChatGPT's ad copy first, and 50 percent of the participants were presented with the freelancer's ad copy first. Before assessing each copy, participants were told whether the headline was written by AI or a human. The validated scales i.e., statements by Smith et al. for measuring (ad) divergence and relevance were adopted i.e., adapted (Smith et al., 2007). However, only one statement per first-order factor (e.g., originality as a first-order factor of divergence) was used to keep the survey as short as possible and to increase the response rates. Furthermore, negatively worded statements (e.g., *"I do NOT care about this product/service"*) were excluded in order to avoid confusion among participants (Bhattacherjee, 2012, p. 76; Smith et al., 2007, p. 831).

Participants who failed an attention test (e.g., could not remember the advertised product) were excluded from the study. The collected data was analysed using SPSS.

The examined advertising copies as well as the questionnaire introduction and the operationalisation table can be found below.

Questionnaire Introduction

Dear Participant,

Thank you for taking the time to take part in this questionnaire survey. Your information is valuable and will make an important contribution to scientific research. Before we begin, I would like to assure you that your answers will be treated confidentially and explain how your data will be used.

Confidentiality: *Please be assured that all information provided in this survey will be kept strictly confidential. Your responses will be kept strictly confidential. Your personal information will be separated from the survey responses so that your individual responses remain anonymous.*

Use of data: *The data collected in this survey will be used for scientific research purposes only. It will be analysed collectively to identify trends, patterns and insights that contribute to a deeper understanding of the topic under study. Individual responses will not be published or made identifiable in any way.*

Reporting of results: *The results of this survey are published in an aggregated format. This means that your individual responses will be combined with those of other participants to ensure that no individual can be identified from the data reported. The aim is to provide an overview of the results, highlight important trends and provide meaningful insights. Your participation will help to produce a comprehensive report that can be useful to the academic community and potentially inform future research and practice.*

Voluntary participation: *Your participation in this survey is entirely voluntary and you have the right to stop participating at any time without giving any reason. There will be no negative consequences if you decide not to participate in the survey or to stop participating during the survey.*

Your contribution is invaluable and will add to the body of knowledge in this area. If at any time you have any questions or concerns, please do not hesitate to contact me.

Yours sincerely,
Okan Karakas

Table 3 - Operationalisation Table

Number	Description	Question/ Statement	Scale expression	H1/2	Citation
1	Exclusionary question: if "no" then exclusion	Do you like ice cream?	• Yes • No	-	-
2	Exclusionary question: if "no" then exclusion	Do you know AI?	• Yes • No	-	-
3	Demographic Data	What is your age?	Open question	-	-
4	Demographic Data	What is your gender?	• Male • Female • Divers / non-binary	-	-
5	Demographic Data	Where is your home located?	• North America/Central America • South America • Europe • Africa • Asia • Australia • Caribbean Islands • Pacific Islands • Prefer not to say	-	-
6	Attitude towards AI	I think AI can be creative on its own.	• completely right • quite right • undecided • quite wrong • completely wrong	H2	Adopted from (Hong et al., 2021, p. 7)
7	Attitude towards AI	I believe AI can make something new by itself.	• completely right • quite right • undecided • quite wrong • completely wrong	H2	Adopted from (Hong et al., 2021, p. 7)

Table 3 - Continued

8	Attitude towards AI	Products developed by AI should be respected as creative works.	• completely right • quite right • undecided • quite wrong • completely wrong	H2	Adopted from (Hong et al., 2021, p. 7)
Presentation of BOTH advertising copies					
9	Awareness Check	What was the text just presented about?	• Toothpaste • Ice cream • Smoothie Maker • Earbuds • Bed Sheets	-	-
Assessment of the written advertising copy (written by human OR ChatGPT)					
10	Measure of divergence factor “Originality”	The text is "out of the ordinary".	• completely right • quite right • undecided • quite wrong • completely wrong	H1	Adapted from (Smith et al., 2007, pp. 830–832)
11	Measure of divergence factor “Flexibility”	The text contains different ideas.	• completely right • quite right • undecided • quite wrong • completely wrong	H1	Adapted from (Smith et al., 2007, pp. 830–832)
12	Measure of divergence factor “Synthesis”	The text contains unusual connections.	• completely right • quite right • undecided • quite wrong • completely wrong	H1	Adapted from (Smith et al., 2007, pp. 830–832)
13	Measure of divergence factor “Elaboration”	The text contains numerous details.	• completely right • quite right • undecided • quite wrong • completely wrong	H1	Adapted from (Smith et al., 2007, pp. 830–832)

Table 3 - Continued

14	Measure of divergence factor “Artistic Value”	The text makes ideas come to life verbally.	• completely right • quite right • undecided • quite wrong • completely wrong	H1	Adapted from (Smith et al., 2007, pp. 830–832)
15	Measure of relevance factor “Ad-to-Consumer Relevance”	The text is appropriate to me.	• completely right • quite right • undecided • quite wrong • completely wrong	H1	Adapted from (Smith et al., 2007, pp. 830–832)
16	Measure of relevance factor “Brand-to-Consumer Relevance”	The product is appropriate to me.	• completely right • quite right • undecided • quite wrong • completely wrong	H1	Adapted from (Smith et al., 2007, pp. 830–832)
Assessment of the written advertising copy (written by human OR ChatGPT)					
Repetition of questions i.e., statements 10 to 16					

Note. This table was created by the author of this paper.

Examined copies i.e., headlines

ChatGPT:

Indulge in Blissful Delights: Unleash Your Sweetest Desires with Our Heavenly Ice Cream Creations!

Human Copywriter:

Release your inner child! Melt away your worries with our handcrafted ice cream creations!

Chapter 4: Empirical study

The survey comprised 131 individuals who voluntarily participated. To ensure data quality, a screening procedure was conducted, resulting in the exclusion of a subset of the participants who failed to meet the required conditions outlined by two exclusionary questions at the beginning of the survey and an awareness i.e., attention check. One participant was excluded due to his age (under 18 years). After this comprehensive screening procedure, the final data set included 75 participants *(n=75)* who successfully completed the entire survey and were considered for further analysis.

Based on the sample data, the average age of the participants was *28.2 years (SD = .86)*, while the age range spanned from 18 to 55 years. Regarding gender distribution, 56 percent of respondents identified as male, 40 percent identified as female, and 4 percent as divers/non-binary.

With the descriptive statistics providing a general overview of the collected data, reliability tests were conducted to estimate the internal consistency (i.e., item interrelatedness) of the measurement scales used. Cronbach's alpha was calculated from the mean values of the individual items. Although both studies from which the theoretical model was derived, see (Hong et al., 2021) for the construct "Attitude towards (creative) AI, and (Smith et al., 2007) for the constructs "Divergence" and "Relevance", conducted a reliability analysis of their measurement scales using Cronbach's alpha (α), in this study another reliability analysis using Cronbach's alpha (α) was conducted for several reasons. As stated by Streiner, reliability scores vary among different sample populations (Streiner, 2003, p. 101). Since there is most likely heterogeneity between the group examined in this study and those in the two studies just mentioned, it is also likely that the reliability values differ. Furthermore, Cronbach's alpha increases proportional to the length of a scale (Streiner, 2003, pp. 100–101).

More specifically, the relationship between alpha and the number of (questionnaire) items is curvilinear, i.e., α initially increases and then becomes smaller or stabilises as the number of items increases (Vaske et al., 2017, p. 165). As the number of scales measuring "Divergence" and "Relevance" was reduced (to increase the response rate), and the consistency (i.e., item interrelatedness) may vary across different survey setups (Vaske et al., 2017, p. 165), conducting a reliability test (i.e., estimation) before further analysis is considered crucial. In addition, Smith et al. estimated the internal consistency of the measurement scales without any relation to AI. Therefore, in this study Cronbach's alpha was calculated for the constructs "Attitude towards (creative) AI", "Human Divergence", "Human Relevance", "AI Divergence", and "AI Relevance". The table below provides an overview of the estimated values.

Table 4 - Reliability tests using Cronbach's alpha (α)

Construct	**Description**	**Cronbach's alpha (α)**	**Number of Items**
Attitude towards (creative) AI	Internal consistency of attitude scales	*.622*	3
Human Divergence	Internal consistency of divergence scales for human-written copy	*.603*	5
Human Relevance	Internal consistency of relevance scales for human-written copy	*.743*	2
AI Divergence	Internal consistency of divergence scales for AI-generated copy	*.550*	5
AI Relevance	Internal consistency of relevance scales for AI-generated copy	*.770*	2

Note. This table was created by the author of this paper.

As *"a high value of α is a prerequisite for internal consistency, but does not guarantee it"* (Streiner, 2003, p. 102), and usually has a value between 0 and 1 (negative values are also possible), the question arises whether the estimated α-values are "acceptable". While the size of alpha of both “Human Relevance” and “AI Relevance” are quite high, the same cannot be stated about the values for “Attitude towards (creative) AI”, “Human Divergence”, and “AI Divergence”. However, as previously mentioned, a low(er) alpha was expected for “Divergence” and “Relevance” due to the low number of items (compared to the study by Smith et al.). Moreover, there seems to be no agreement in research on "acceptable" alpha values. As cited by Streiner and Cho & Kim, Nunnally (1967) initially advised researchers to aim for alpha values around .50 to .60 during the early stages of research, .80 for fundamental research instruments, and at least .90 for clinical applications, aiming for an ideal of .95. However, Nunnally revised the starting point to .70 in later iterations of his book (Cho & Kim, 2015, p. 217; Streiner, 2003, p. 103). On the contrary, Vaske et al., for example, are not as specific and consider alpha values between .65 and .80 acceptable (Vaske et al., 2017, p. 168). Nevertheless, the estimated alpha values in this study are considered acceptable since (a) most values are close to the suggested values for early stages of research by Nunnally, and (b) fewer items per construct (for “Divergence” and “Relevance”) were analysed, consequently resulting in lower alpha values. Lastly, it should be mentioned that the deletion of lower correlating items (analysis of only higher correlating items), did not result in a substantial increase of the alpha value (Cho & Kim, 2015, p. 217; Vaske et al., 2017, p. 171).

Testing of hypothesis 1 (H1)

After conducting a reliability analysis of the scales, a one-way repeated measures multivariate analysis of variance (i.e., one-way repeated measures MANOVA) was carried out to investigate Hypothesis 1 (H1). A multivariate approach was chosen for several reasons. First to be mentioned is that both second-order composite latent factors “Divergence” and

"Relevance" are considered dependent variables, and when taken together (two dependent variables separately examined in one method), form a meaningful group of the variable under investigation ("Perceived Advertising Creativity") (Pituch & Stevens, 2016, p. 142). Second, other than univariate tests (e.g., ANOVA), a MANOVA automatically corrects for alpha error inflation, and thus increases statistical power. Furthermore, multivariate tests consider the correlations among the variables, which are ignored in univariate tests (Pituch & Stevens, 2016, p. 143). Since the measurement was done within one group and among individuals without interaction (i.e., independence of observations), and the dependent variables are measured at the interval level, while the independent variables (AI and human copy) are categorical, the most crucial assumptions for a MANOVA are given (*One-Way Repeated Measures MANOVA in SPSS Statistics*, n.d.; Pituch & Stevens, 2016, p. 220). In addition, as each participant had to rate "Divergence" and "Relevance" for human-written and AI-generated copy (i.e., took part in all treatments), a repeated measure design is deemed appropriate (O'Brien & Kaiser, 1985, p. 316). For the purpose of testing the hypotheses, a significance level of *5%* ($\vartheta = .05$) has been assumed in this study. The results suggest that there is a significant difference in the dependent variables "Divergence" and "Relevance" between AI-generated copy and human-written copy (*Divergence:* $p = .002 < .05$, $F = 10.49$ | *Relevance:* $p = .034 < .05$, $F = 4.65$). Since there is only one pairwise comparison (AI-generated vs human-written copy), no post hoc testing was conducted, and differences were determined by comparing the means (*One-Way Repeated Measures MANOVA in SPSS Statistics*, n.d.). Human-written copy received higher ratings for the "Relevance" variable ($Human_R$: $M = 3.93$, $SD = .86$ | AI_R: $M = 3.65$, $SD = .97$), while AI-generated copy achieved higher ratings for the variable "Divergence" ($Human_D$: $M = 3.18$, $SD = .69$ | AI_D: $M = 3.49$, $SD = .62$). Therefore, H1 can be (partially) rejected as there is insufficient statistical evidence that human intervention has a positive impact on creative output (i.e., the evaluation of that output). Lastly, it should be mentioned that both dependent variables, "Divergence" and "Relevance," exhibit effect sizes of moderate magnitude *(partial eta squared = .124 and .059)* (Norouzian & Plonsky, 2018, p. 267).

Testing of hypothesis 2 (H2)

Following the examination of hypothesis 1, a bivariate correlation analysis was conducted to analyse hypothesis 2 (H2). Thus, the strength of the relationship between **(a)** "attitude towards (creative) AI" and "AI divergence" and **(b)** "attitude towards (creative) AI" and "AI relevance" was examined. Given that the scales are interval scales and the relationship between those variables was examined, a Pearson product moment correlation (hereinafter, correlation) was performed. As with most statistical analyses, the significance level was again set at *5%* ($\vartheta = .05$) (Bhattacherjee, 2012, p. 125). However, the author is well aware that bivariate correlations, as the name suggests, examine the relationship between two variables (Bhattacherjee, 2012, p. 122). Since two correlations were performed simultaneously, a Bonferroni-adjusted procedure was used to avoid inflating the alpha error, resulting in a reduced significance level of *.025 (.05/2)* (Wright, 1992, p. 1008). Results suggest that both relationships (a and b) are not significant ($p_a = .039 > .025; p_b = .149 > .025$), and minimal positive correlations are present ($r_a = .204; r_b = .122$). Consequently, there does not seem to be sufficient statistical evidence that there is a positive relationship between the belief that AI can be creative and the rating of its texts. H2 is therefore rejected.

Chapter 5: Concluding remarks

The aim of this study was to investigate the potential impact of AI applications (i.e., LLMs) such as ChatGPT on the profession of professional copywriting. More specifically, the study wanted to find out whether LLMs pose a threat to the traditional role of copywriters or whether they are simply valuable tools that support the creative process. By comparing AI-generated and human-written advertising copy, the study aimed to shed light on the differences between these two sources in terms of creativity (H1). In addition, it investigated whether people's attitudes towards creative AI, i.e., their beliefs about AI's ability to be creative, has an impact on the assessment of AI-generated advertising copy (H2).

The results of testing the first hypothesis showed significant differences in consumer perceptions between AI-generated and human-written advertising copy. The AI-generated text was found to excel in terms of being divergent, a quality that contributes significantly to the creativity of advertising. Hence, the AI-generated copy received higher ratings for their uniqueness and novel ideas. On the other hand, the human-written copy was perceived as more relevant and valuable by the consumers under research. This result suggests that while ChatGPT can be innovative and bring new ideas to the table, human copywriters are still better at writing content that speaks to target audiences on a deeper level.

Interestingly, the study found that attitudes towards creative AI did not have a significant impact on the evaluation of AI-generated ad copy. This could mean that consumers' perceptions of AI-generated content are based on the actual quality of the output rather than their preconceived beliefs about the creative capabilities of AI.

Despite the positively perceived contributions of ChatGPT to the creative process, the question remains whether this technology will eventually replace copywriters altogether. The data collected in this study was not sufficient to make definitive predictions about the impact of ChatGPT on the professional job of copywriters. While ChatGPT has proven its value as a helpful copywriting tool, its potential to completely replace human copywriters remains uncertain.

In addition, the study acknowledged that AI technology and applications, including chatbots such as ChatGPT, will continue to improve their efficiency and effectiveness. With the advances in AI algorithms, it is possible that chatbots could not only outperform human copywriters in terms of "divergence", but also catch up with or even surpass them in terms of relevance, i.e., meaningfulness. This could pose greater challenges to human copywriters in the future and push the industry to adapt and evolve with technological advances.

It is important to recognise that the contributions of this study go beyond its specific focus on copywriting. By exploring the relatively under-researched topic of content marketing in the context of AI, the study provides valuable insights into the opinions and perceptions of both researchers and the general population regarding ChatGPT and its potential implications for marketers.

In summary, ChatGPT and similar AI technologies, while promising as helpful tools in the creative process, cannot yet fully replace the human copywriter. The study highlights the importance of recognising the strengths and limitations of both AI-generated and human copywriting, as well as the need for ongoing research and adaptation in response to rapid advances in AI technology. By understanding the symbiotic relationship between AI and human creativity, the field of professional copywriting can potentially harness the capabilities of AI while preserving the unique touch that human writers bring to the table.

Chapter 6: Limitations and future research

Every scientific study, including this one, has its inherent limitations that should be acknowledged for further discussion. Therefore, this final chapter addresses the limitations of the study and discusses possible avenues for further research.

First of all, it should be highlighted that a convenience sample was used, which may limit the generalizability of the findings. While internet-based sampling has certainly its advantages, surveys show that internet users are not truly representative of the global population, even in 2023. For instance, the income level of a country is a decisive factor for internet access. While more than 90 percent of people in high-income countries have access to the internet, this percentage accumulates to only 26 percent in low-income countries (*Internet and Social Media Users in the World 2023*, 2023). Therefore, future research should employ alternative (probabilistic) sampling methods and/or examine the effects investigated in this study in specific customer segments to increase the representativeness of the findings and search for variations in the analysed variables across different segments. Furthermore, the relatively small sample size and the use of predominantly students and/or younger age individuals may limit the statistical power of the results presented. Therefore, it is recommended to use larger and more diverse samples in future research avenues to increase statistical power.

The limited focus of the study on a specific text format (headlines) and product category (ice cream) may limit the scope of the findings as well. In order to gain a more comprehensive understanding of the phenomenon, future research should examine different copywriting formats and different product categories, including low-involvement products. In addition, combining different content formats, such as text and images, can provide valuable insights and help explore potential interaction effects between different variables i.e., scenarios.

The study's limitation to English texts neglects the potential influence of different languages and cultures on the individual evaluation of advertising creativity. To address this limitation, future studies should conduct cross-linguistic analyses to determine how linguistic and cultural factors may influence the observed effects. The disclosure of the respective text sources in the survey might have led to a bias in the participants' responses. Future studies could mitigate this bias by using an experimental design that hides the source of the texts and thus reduces the risk of response bias.

Despite efforts to minimise bias, there may still be unknown variables or contextual factors that may have influenced the results. Researchers should be careful when interpreting the results and consider possible confounding factors in their analysis.

Quantitative questionnaire surveys, while suitable for measuring beliefs, attitudes and preferences, may not capture participants' in-depth perceptions and experiences. The inclusion of qualitative data collection methods can lead to a more sophisticated understanding of participants' views and enrich the research findings.

In summary, it is crucial for the accuracy and reliability of the study's findings to acknowledge and address these limitations. Future research should seek to build on these limitations by using larger and more diverse samples, investigating different content formats and product categories, explore cross-linguistic effects, taking into account possible confounding factors, and incorporating both quantitative and qualitative data collection methods. In this way, researchers can improve the validity, reliability and applicability of their findings and provide more robust insights to both the scientific community and the commercial world. As AI and creative technologies develop, comprehensive and well-designed studies are essential to effectively navigate the impact on the field of professional copywriting.

References

Achar, C., So, J., Agrawal, N., & Duhachek, A. (2016). What we feel and why we buy: The influence of emotions on consumer decision-making. *Current Opinion in Psychology*, *10*, 166–170. https://doi.org/10.1016/j.copsyc.2016.01.009

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, *2*. https://doi.org/10.1016/j.mlwa.2020.100006

Agudo, U., Arrese, M., Liberal, K. G., & Matute, H. (2022). Assessing Emotion and Sensitivity of AI Artwork. *Frontiers in Psychology*, *13*, 879088. https://doi.org/10.3389/fpsyg.2022.879088

Albers, S., Klapper, D., Konradt, U., Walter, A., & Wolf, J. (Eds.). (2009). *Methodik der empirischen Forschung*. Gabler Verlag. https://doi.org/10.1007/978-3-322-96406-9

Ananthakrishnan, R., & Arunachalam, D. T. (2022). *Comparison Of Consumers Perception Between Human Generated And Ai Aided Brand Content*. *19*(2).

Arif, T. B., Munaf, U., & Ul-Haque, I. (2023). The future of medical education and research: Is ChatGPT a blessing or blight in disguise? *Medical Education Online*, *28*(1), 2181052. https://doi.org/10.1080/10872981.2023.2181052

Baumgarth, C., & Kirkby, A. (2022). KI und Markenstimme—Optionen, Erkenntnisse und Agenden. *Transfer: Zeitschrift Für Kommunikation & Markenmanagement*, *68*(3), 32–41.

Bhattacherjee, A. (2012). *Social science research: Principles, methods, and practices* (Second edition). Anol Bhattacherjee.

Bhirud, N., Tatale, S., Randive, S., & Nahar, S. (2019, August 1). *A Literature Review On Chatbots In Healthcare Domain*.

Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C.-H. (2019). Artificial Intelligence and Journalism. *Journalism & Mass Communication Quarterly*, *96*(3), 673–695. https://doi.org/10.1177/1077699019859901

Bünte, C. (2021). Hype und Wirklichkeit: Über Image, Einsatz und Nutzen von KI in der Werbung. *Transfer: Zeitschrift Für Kommunikation & Markenmanagement*, *67*(2), 32–37.

Cai, Y. (Ed.). (2013). *3D Immersive and Interactive Learning*. Springer. https://doi.org/10.1007/978-981-4021-90-6

Campbell, C., Plangger, K., Sands, S., & Kietzmann, J. (2022). Preparing for an Era of Deepfakes and AI-Generated Ads: A Framework for Understanding Responses to Manipulated Advertising. *Journal of Advertising*, *51*(1), 22–38. https://doi.org/10.1080/00913367.2021.1909515

Carpenter, K. A., & Altman, R. B. (2023). Using GPT-3 to Build a Lexicon of Drugs of Abuse Synonyms for Social Media Pharmacovigilance. *Biomolecules*, *13*(2), 387. https://doi.org/10.3390/biom13020387

Castro, D., & New, J. (2016). *The Promise of Artificial Intelligence*. 48.

ChatGPT. (2023). https://chat.openai.com

ChatGPT, & Affairs, J. of I. (2022). OPENAI'S CHATGPT AND THE PROSPECT OF LIMITLESS INFORMATION: A Conversation with ChatGPT. *Journal of International Affairs*, *75*(1), 379–386.

Chintalapati, S., & Pandey, S. K. (2022). Artificial intelligence in marketing: A systematic literature review. *International Journal of Market Research*, *64*(1), 38–68. https://doi.org/10.1177/14707853211018428

Cho, E., & Kim, S. (2015). Cronbach's Coefficient Alpha: Well Known but Poorly Understood. *Organizational Research Methods*, *18*(2), 207–230. https://doi.org/10.1177/1094428114555994

Chuang, H.-M., & Cheng, D.-W. (2022). Conversational AI over Military Scenarios Using Intent Detection and Response Generation. *Applied Sciences*, *12*(5), 2494. https://doi.org/10.3390/app12052494

Coffin, J. (2022). Asking Questions of AI Advertising: A Maieutic Approach. *Journal of Advertising*, *51*(5), 608–623. https://doi.org/10.1080/00913367.2022.2111728

Dale, R. (2021). GPT-3: What's it good for? *Natural Language Engineering*, *27*(1), 113–118. https://doi.org/10.1017/S1351324920000601

DALL·E 2. (2023, April 10). https://openai.com/product/dall-e-2

Darlington, K. (2023). ChatGPT Is A Ground-Breaking Application Of AI That We Will All Be Hearing About In The Near Future. *IQ: The RIM Quarterly*, *39*(1), 49–51.

de-Lima-Santos, M.-F., & Ceron, W. (2021). Artificial Intelligence in News Media: Current Perceptions and Future Outlook. *Journalism and Media*, *3*(1), 13–26. https://doi.org/10.3390/journalmedia3010002

Deng, S., Tan, C.-W., Wang, W., & Pan, Y. (2019). Smart Generation System of Personalized Advertising Copy and Its Application to Advertising Practice and Research. *Journal of Advertising*, *48*(4), 356–365. https://doi.org/10.1080/00913367.2019.1652121

Dong, C., Li, Y., Gong, H., Chen, M., Li, J., Shen, Y., & Yang, M. (2023). A Survey of Natural Language Generation. *ACM Computing Surveys*, *55*(8), 1–38. https://doi.org/10.1145/3554727

El-Murad, J., & West, D. C. (2004). The Definition and Measurement of Creativity: What Do We Know? *Journal of Advertising Research*, *44*(2), 188–201. https://doi.org/10.1017/S0021849904040097

Farnworth, D. (2015). What Is Content Marketing? *Copyblogger*. https://copyblogger.com/content-marketing-codex/

Fatima, N., Imran, A. S., Kastrati, Z., Daudpota, S. M., & Soomro, A. (2022). A Systematic Literature Review on Text Generation Using Deep Neural Network Models. *IEEE Access*, *10*, 53490–53503. https://doi.org/10.1109/ACCESS.2022.3174108

Floridi, L. (2023). AI as Agency Without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models. *Philosophy & Technology*, *36*(1), 15, s13347-023-00621-y. https://doi.org/10.1007/s13347-023-00621-y

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds and Machines*, *30*(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1

Gartner. (2015, October 6). *Top 10 Predictions herald what it means to be human in a digital world*. Gartner Predicts Our Digital Future. https://www.gartner.com/smarterwithgartner/gartner-predicts-our-digital-future

Gatt, A., & Krahmer, E. (2018). Survey of the State of the Art in Natural Language Generation: Core tasks, applications and evaluation. *Journal of Artificial Intelligence Research*, *61*, 65–170. https://doi.org/10.1613/jair.5477

Ge, J., & Lai, J. C. (2023). Artificial intelligence-based text generators in hepatology: ChatGPT is just the beginning. *Hepatology Communications*, *7*(4). https://doi.org/10.1097/HC9.0000000000000097

Global search engine desktop market share 2023. (n.d.). Statista. Retrieved 12 April 2023, from https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

GPT-4. (n.d.). Retrieved 12 April 2023, from https://openai.com/research/gpt-4

Hacker, P. (2023). Understanding and Regulating ChatGPT, and Other Large Generative AI Models. *Verfassungsblog: On Matters Constitutional*. https://doi.org/10.17176/20230120-220055-0

Hauck, M., & Pagel, S. (2020). AI Media Technology Landscape—Systematisierungsinstrument für den Einsatz von KI in Medienunternehmen. *Innovation in der Medienproduktion und -distribution - Proceedings der Jahrestagung der Fachgruppe Medienökonomie der DGPUK 2019*, Köln. https://doi.org/10.21241/SSOAR.68092

Holliman, G., & Rowley, J. (2014). Business to business digital content marketing: Marketers' perceptions of best practice. *Journal of Research in Interactive Marketing*, *8*(4), 269–293. https://doi.org/10.1108/JRIM-02-2014-0013

Hong, J. W., Peng, Q., & Williams, D. (2021). Are you ready for artificial Mozart and Skrillex? An experiment testing expectancy violation theory and AI music. *New Media & Society*, *23*(7), 1920–1935. https://doi.org/10.1177/1461444820925798

Hussain, S., Ameri Sianaki, O., & Ababneh, N. (2019). A Survey on Conversational Agents/Chatbots Classification and Design Techniques. In L. Barolli, M. Takizawa, F. Xhafa, & T. Enokido (Eds.), *Web, Artificial Intelligence and Network Applications* (Vol. 927, pp. 946–956). Springer International Publishing. https://doi.org/10.1007/978-3-030-15035-8_93

Internet and social media users in the world 2023. (2023). Statista. https://www.statista.com/statistics/617136/digital-population-worldwide/

Januszewicz, J., Jarosz, M., Smoliński, P., & Winiarski, J. (2022). IMPORTANCE OF EMOTIONS IN ADVERTISING: ASSESSMENT OF DIFFERENCES IN EMOTION LEVELS BETWEEN ADVERTISING TEXT CREATED BY COPYWRITERS AND AI IN THE PHARMACEUTICAL INDUSTRY. *Contemporary Economy*, *14*(2/S). https://czasopisma.bg.ug.edu.pl/index.php/wg/article/view/7117

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Kemp, E., Bui, M., & Chapa, S. (2012). The role of advertising in consumer emotion management. *International Journal of Advertising*, *31*(2), 339–353. https://doi.org/10.2501/IJA-31-2-339-353

Köse, U., & Sert, S. (2016). *Intelligent Content Marketing with Artificial Intelligence*. https://www.academia.edu/28685210/Intelligent_Content_Marketing_with_Artificial_Intelligence

Kotenidis, E., & Veglis, A. (2021). Algorithmic Journalism—Current Applications and Future Perspectives. *Journalism and Media*, *2*(2), 244–257. https://doi.org/10.3390/journalmedia2020014

Kover, A. J. (1995). Copywriters' Implicit Theories of Communication: An Exploration. *Journal of Consumer Research*, *21*(4), 596–611. https://doi.org/10.1086/209421

Kover, A. J. (1996). Why Copywriters Don't Like Advertising Research—And What Kind of Research Might They Accept. *Journal of Advertising Research*, *36*(2), RC-8.

Kover, A. J. (2016). Advertising Creativity: Some Open Questions. *Journal of Advertising Research*, *56*(3), 235. https://doi.org/10.2501/JAR-2016-033

Kucherbaev, P., Bozzon, A., & Houben, G.-J. (2018). Human-Aided Bots. *IEEE Internet Computing*, *22*(6), 36–43. https://doi.org/10.1109/MIC.2018.252095348

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., Leon, L. D., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, *2*(2), e0000198. https://doi.org/10.1371/journal.pdig.0000198

Kursh, S. (2021). An Introduction to the 'How To' for AI and Machine Learning. *Business Education Innovation Journal*, *13*(2), 14–23.

Leo Kumar, S. p. (2019). Knowledge-based expert system in manufacturing planning: State-of-the-art review. *International Journal of Production Research*, *57*(15/16), 4766–4790. https://doi.org/10.1080/00207543.2018.1424372

Lin, C.-C., Huang, A. Y. Q., & Yang, S. J. H. (2023). A Review of AI-Driven Conversational Chatbots Implementation Methodologies and Challenges (1999–2022). *Sustainability*, *15*(5), 4012. https://doi.org/10.3390/su15054012

Linwan Wu, & Jing Wen, T. (2021). Understanding AI Advertising From the Consumer Perspective: What Factors Determine Consumer Appreciation of AI-Created Advertisements? *Journal of Advertising Research*, *61*(2), 133–146. https://doi.org/10.2501/JAR-2021-004

Lund, B., Ting, W., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a New Academic Reality: AI-Written Research Papers and the Ethics of the Large Language Models in Scholarly Publishing. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4389887

Maslen, A. (2010). *The copywriting sourcebook: How to write better copy, faster--for everything from ads to websites*. Marshall Cavendish Business.

McCarthy, J. (1998). *WHAT IS ARTIFICIAL INTELLIGENCE?* 12.

Melin, M., Back, A., Sodergard, C., Munezero, M. D., Leppanen, L. J., & Toivonen, H. (2018). No Landslide for the Human Journalist—An Empirical Study of Computer-Generated Election News in Finland. *IEEE Access*, *6*, 43356–43367. https://doi.org/10.1109/ACCESS.2018.2861987

Motger, Q., Franch, X., & Marco, J. (2022). Software-Based Dialogue Systems: Survey, Taxonomy, and Challenges. *ACM Computing Surveys*, *55*(5), 91:1-91:42. https://doi.org/10.1145/3527450

Noain-Sánchez, A. (2022). Addressing the Impact of Artificial Intelligence on Journalism: The perception of experts, journalists and academics. *Communication & Society*, *35*(3), 105–121. https://doi.org/10.15581/003.35.3.105-121

Norouzian, R., & Plonsky, L. (2018). Eta- and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research*, *34*(2), 257–271. https://doi.org/10.1177/0267658316684904

O'Brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: An extensive primer. *Psychological Bulletin*, *97*(2), 316–333. https://doi.org/10.1037/0033-2909.97.2.316

One-way repeated measures MANOVA in SPSS Statistics. (n.d.). Retrieved 17 July 2023, from https://statistics.laerd.com/spss-tutorials/one-way-repeated-measures-manova-using-spss-statistics.php

OpenAI. (2023a). *GPT-4 Technical Report* (arXiv:2303.08774). arXiv. http://arxiv.org/abs/2303.08774

OpenAI. (2023b, April 11). *Planning for AGI and beyond*. Planning for AGI and Beyond. https://openai.com/blog/planning-for-agi-and-beyond

OpenAI Platform. (n.d.). Retrieved 20 July 2023, from https://platform.openai.com

Osmanovic-Thunström, A., & Steingrimsson, S. (2023). Does GPT-3 qualify as a co-author of a scientific paper publishable in peer-review journals according to the ICMJE criteria? A case study. *Discover Artificial Intelligence*, *3*(1), 12. https://doi.org/w

Overgoor, G., Chica, M., Rand, W., & Weishampel, A. (2019). Letting the Computers Take Over: Using AI to Solve Marketing Problems. *California Management Review*, *61*(4), 156–185. https://doi.org/10.1177/0008125619859318

Pais, S., Cordeiro, J., & Jamil, M. L. (2022). NLP-based platform as a service: A brief review. *Journal of Big Data*, *9*(1), 54. https://doi.org/10.1186/s40537-022-00603-5

Pereira, J., & Díaz, O. (2018). *Chatbot Dimensions that Matter: Lessons from the Trenches* (pp. 129–135). https://doi.org/10.1007/978-3-319-91662-0_9

Pérez-Soler, S., Juarez-Puerta, S., Guerra, E., & Lara, J. (2021). Choosing a Chatbot Development Tool. *IEEE Software*, *PP*. https://doi.org/10.1109/MS.2020.3030198

Pituch, K. A., & Stevens, J. (2016). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS* (6th edition). Routledge/Taylor & Francis Group.

Qin, X., & Jiang, Z. (2019). The Impact of AI on the Advertising Process: The Chinese Experience. *Journal of Advertising*, *48*(4), 338–346. https://doi.org/10.1080/00913367.2019.1652122

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving Language Understanding by Generative Pre-Training*. https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners. *February 14, 2019*.

Ramesh, A., & Chawla, V. (2022). Chatbots in Marketing: A Literature Review Using Morphological and Co-Occurrence Analyses. *Journal of Interactive Marketing*, *57*(3), 472–496. https://doi.org/10.1177/10949968221095549

Rana, S. (2023). AI and GPT for Management Scholars and Practitioners: Guidelines and Implications. *FIIB Business Review*, *12*(1), 7–9. https://doi.org/10.1177/23197145231161408

Renn, O. (2023). How sustainable is the digital world? *GAIA: Ecological Perspectives for Science & Society*, *32*, 1–1. https://doi.org/10.14512/gaia.32.S1.1

Rozado, D. (2023). The Political Biases of ChatGPT. *Social Sciences*, *12*(3), 148. https://doi.org/10.3390/socsci12030148

Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (Fourth edition). Pearson.

Sai, A. B., Mohankumar, A. K., & Khapra, M. M. (2022). A Survey of Evaluation Metrics Used for NLG Systems. *ACM Computing Surveys*, *55*(2), 26:1-26:39. https://doi.org/10.1145/3485766

Salido, M. A. (2010). Introduction to planning, scheduling and constraint satisfaction. *Journal of Intelligent Manufacturing*, *21*(1), 1–4. https://doi.org/10.1007/s10845-008-0188-z

Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing? *Critical Care*, *27*(1), 75. https://doi.org/10.1186/s13054-023-04380-2

Samoili, S., Lopez, C. M., Delipetrev, B., Martinez-Plumed, F., Gomez, G. E., & De, P. G. (2021). AI Watch. Defining Artificial Intelligence 2.0. *JRC Publications Repository*. https://doi.org/10.2760/019901

Sarstedt, M., Bengart, P., Shaltoni, A. M., & Lehmann, S. (2018). The use of sampling methods in advertising research: A gap between theory and practice. *International Journal of Advertising*, *37*(4), 650–663. https://doi.org/10.1080/02650487.2017.1348329

Serdouk, A., & Bessam, A. C. (2023). Bots in Newsrooms: What Future for Human Journalists? *Media Watch*, *14*(1), 100–115. https://doi.org/10.1177/09760911221130816

Shankar, V. (2018). How Artificial Intelligence (AI) is Reshaping Retailing. *Journal of Retailing*, *94*(4), vi–xi. https://doi.org/10.1016/S0022-4359(18)30076-9

Skibba, R. (2020). Media Enhanced by Artificial Intelligence: Can We Believe Anything Anymore? *Engineering*, *6*(7), 723–724. https://doi.org/10.1016/j.eng.2020.05.011

Smith, R. E., MacKenzie, S. B., Xiaojing Yang, Buchholz, L. M., & Darley, W. K. (2007). Modeling the Determinants and Effects of Creativity in Advertising. *Marketing Science*, *26*(6), 819–833. https://doi.org/10.1287/mksc.1070.0272

Smith, R. E., & Yang, X. (2004). Toward a General Theory of Creativity in Advertising: Examining the Role of Divergence. *Marketing Theory*, *4*(1–2), 31–58. https://doi.org/10.1177/1470593104044086

Streiner, D. L. (2003). Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment*, *80*(1), 99–103. https://doi.org/10.1207/S15327752JPA8001_18

Sumitha, S. (2022). A Comparative Analysis of Artificial Intelligence in Marketing and Traditional Marketing. *International Journal of Business Analytics & Intelligence (IJBAI)*, *10*(1), 16–21.

Sundar, S. S. (2020). Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAII). *Journal of Computer-Mediated Communication*, *25*(1), 74–88. https://doi.org/10.1093/jcmc/zmz026

Szeliski, R. (2022). *Computer vision: Algorithms and applications, 2nd ed.* (Second edition). Springer. https://szeliski.org/Book

Taecharungroj, V. (2023). "What Can ChatGPT Do?" Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. *Big Data and Cognitive Computing*, *7*(1), 35. https://doi.org/10.3390/bdcc7010035

Teubner, T., Flath, C. M., Weinhardt, C., van der Aalst, W., & Hinz, O. (2023). Welcome to the Era of ChatGPT et al.: The Prospects of Large Language Models. *Business & Information Systems Engineering*, *65*(2), 95–101. https://doi.org/10.1007/s12599-023-00795-x

Till, B. D., & Baack, D. W. (2005). RECALL AND PERSUASION: Does Creative Advertising Matter? *Journal of Advertising*, *34*(3), 47–57. https://doi.org/10.1080/00913367.2005.10639201

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

Totschnig, W. (2020). Fully Autonomous AI. *Science and Engineering Ethics*, *26*(5), 2473–2485. https://doi.org/10.1007/s11948-020-00243-z

Turing, A. M. (1950). I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*, *LIX*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

Vakratsas, D., & Ambler, T. (1999). How Advertising Works: What Do We Really Know? *Journal of Marketing*, *63*(1), 26–43. https://doi.org/10.2307/1251999

Van Noort, G., Himelboim, I., Martin, J., & Collinger, T. (2020). Introducing a Model of Automated Brand-Generated Content in an Era of Computational Advertising. *Journal of Advertising*, *49*(4), 411–427. https://doi.org/10.1080/00913367.2020.1795954

Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking Internal Consistency in Cronbach's Alpha. *Leisure Sciences*, *39*(2), 163–173. https://doi.org/10.1080/01490400.2015.1127189

Vogt, M. (2018). An overview of deep learning techniques. *At - Automatisierungstechnik*, *66*(9), 690–703. https://doi.org/10.1515/auto-2018-0076

Volkmar, G., Reinecke, S., & Fischer, P. M. (2021). Künstliche Intelligenz im Marketing: Möglichkeiten und Herausforderungen: Artificial Intelligence in Marketing: Opportunities and Challenges. *Die Unternehmung / Swiss Journal of Business Research & Practice*, *75*(3), 359–375. https://doi.org/10.5771/0042-059X-2021-3-359

West, D. C., Kover, A. J., & Caruana, A. (2008). Practitioner and Customer Views of Advertising Creativity: Same Concept, Different Meaning? *Journal of Advertising*, *37*(4), 35–46. https://doi.org/10.2753/JOA0091-3367370403

What is Content Marketing? (n.d.). *Content Marketing Institute*. Retrieved 21 May 2023, from https://contentmarketinginstitute.com/what-is-content-marketing/

Wright, S. P. (1992). Adjusted P-Values for Simultaneous Inference. *Biometrics*, *48*(4), 1005. https://doi.org/10.2307/2532694

Yu, W., Zhu, C., Li, Z., Hu, Z., Wang, Q., Ji, H., & Jiang, M. (2022). A Survey of Knowledge-enhanced Text Generation. *ACM Computing Surveys*, *54*(11s), 1–38. https://doi.org/10.1145/3512467

Zhang, M., & Li, J. (2021). A commentary of GPT-3 in MIT Technology Review 2021. *Fundamental Research*, *1*(6), 831–833. https://doi.org/10.1016/j.fmre.2021.11.011

Zhang, X., Zou, Y., Zhang, H., Zhou, J., Diao, S., Chen, J., Ding, Z., He, Z., He, X., Xiao, Y., Long, B., Yu, H., & Wu, L. (2022). Automatic Product Copywriting for E-commerce. *Proceedings of the AAAI Conference on Artificial Intelligence*, *36*(11), 12423–12431. https://doi.org/10.1609/aaai.v36i11.21508

Zhou, J., Ke, P., Qiu, X., Huang, M., & Zhang, J. (2023). ChatGPT: Potential, prospects, and limitations. *Frontiers of Information Technology & Electronic Engineering*. https://doi.org/10.1631/FITEE.2300089

Zimmerman, A. (2023). A Ghostwriter for the Masses: ChatGPT and the Future of Writing. *Annals of Surgical Oncology*, s10434-023-13436–0. https://doi.org/10.1245/s10434-023-13436-0

Zumstein, D., & Hundertmark, S. (2017). Chatbots – An Interactive Technology for Personalized Communication, Transactions and Services. *IADIS International Journal on WWW/Internet*, *15*, 96–109.