Unveiling the multifaceted public interest in ChatGPT

A study on societal implications and operational realities

Abubaker Qutieshat

Oman Dental College, Oman
University of Dundee, United Kingdom

✉ aqutieshat@dundee.ac.uk

Abstract

The present study was aimed at identifying key topics in online discussions about the use of ChatGPT by examining a large dataset extracted from Reddit social media using natural language processing. A corpus of 159,971 posts about ChatGPT were extracted from a custom-made python-coded Reddit content scraper for posts in the r/ChatGPT subreddit discussions. After cleaning the data, the sample was reduced to 119,853 posts which was subjected to cluster analysis using the open-source IRaMuTeQ software to identify main topics based on the cooccurrence of texts. These clusters were named by a panel of social psychology experts (n=3) by reading typical text segments within each cluster. Four thematic clusters emerged, categorized into two main topics: "Society and AI Integration", focusing on ethical concerns (32.1%), and "Operational Aspects and Applications", which delves into technical and practical facets (67.9%). The latter includes clusters like "AI Technical Framework", "Casual AI Interactions", and "Human-AI Etiquette". The Reddit discourse provides a comprehensive understanding of ChatGPT, revealing user priorities like system capabilities and ethical considerations. Notably, the "Human-AI Etiquette" cluster is a new topic less covered in existing literature. The findings underscore the importance of effective prompting for meaningful user engagement with ChatGPT.

Keywords: Artificial Intelligence; ChatGPT; Natural Language Processing; Reddit

1. Introduction

In recent years, conversational agents like OpenAI's ChatGPT have become increasingly integrated into various aspects of daily life, from customer service to personal assistance and beyond. These technologies are not just transforming how we interact with machines but also raising important questions about their broader societal and operational implications. As these artifical intelligence (AI) systems become more

sophisticated and ubiquitous, understanding public perceptions and concerns becomes crucial for both ethical and practical reasons.

While there has been considerable research on the technical aspects of conversational agents, less attention has been paid to the societal and operational dimensions (Pan et al., 2023). Moreover, the existing literature has scarcely touched upon the norms and behaviors governing human interactions with AI systems like ChatGPT. This gap in the literature is particularly noteworthy given the increasing role of these technologies in social and professional settings (Flathmann et al., 2023; Van Pinxteren et al., 2020).

Social media platforms have become fertile grounds for public discourse on emerging technologies (Lian et al., 2023). Reddit, with its 55.79 million daily active users and 1.66 billion monthly active users as of 2023, serves as a significant platform for such discussions. Within Reddit, the r/ChatGPT subreddit stands out with its 3.2 million members, serving as a community hub where users discuss various aspects of ChatGPT, from its technical framework to its societal implications. These discussions offer a unique window into public perceptions and attitudes, providing real-world data that can complement academic research.

In recent years, the exploration of public discourse on platforms like Reddit has attracted substantial academic interest. This surge is driven by the extensive datasets these platforms offer, paving the way for a fusion of computational methods with traditional social sciences. An interdisciplinary approach has been adopted, broadening the horizon of analysis and making it possible to delve deeper into the intricacies of public dialogue (Thukral et al., 2018). This surge is attributable to Reddit's capacity to capture real-time public responses to global events, presenting a dynamic and rich dataset for observation (Proferes et al., 2021). Unlike platforms with character limitations, such as Twitter (now rebranded as X), Reddit's structural design, comprised of self-organized subreddits, facilitates targeted data retrieval and provides a more expansive qualitative and quantitative dataset for researchers. This structure supports the creation of dedicated communities around specific topics, enabling more focused and in-depth discussions (Medvedev et al., 2019).

Reddit's operational mechanics set it apart from other social media platforms in distinctive ways. It enables users to contribute content that becomes instantly accessible for community voting and commenting, thereby cultivating an active environment characterized by rich participation and dynamic dialogue (Jamnik & Lane 2017). This participatory nature is further amplified by Reddit's unique voting system and nested commenting structure, which together foster a more egalitarian and content-centric discussion ecosystem. Unlike platforms where connectivity between users dictates the visibility of content, Reddit places the emphasis on the content itself, facilitated by subreddit-focused discussions (Medvedev et al., 2019).

Anonymity on Reddit plays a pivotal role in shaping interactions. The platform allows users to engage freely in discussions without the constraints and biases often associated with identifiable social media profiles (Jaidka et al 2021). This anonymity can contribute to a more honest and uninhibited exchange of ideas, making Reddit a fertile ground for genuine public discourse. Additionally, Reddit's algorithm, which determines post visibility based on time and voting score (karma), ensures that content remains both relevant and timely (Medvedev et al., 2019). Such features highlight Reddit's capacity to tap into the pulse of diverse online communities, offering researchers a window into unfiltered and wide-ranging public sentiments.

Moreover, the behavioral patterns observed on Reddit, from the distribution of comments on posts to the diversity in user engagement (posting and commenting), shed light on the complex nature of community-driven interactions. This diversity, alongside the significant volume of user-generated content, positions Reddit as a critical resource for understanding the multifaceted dimensions of public discourse, particularly in relation to emerging technological phenomena (Thukral et al., 2018; Pokharel et al 2019).

The burgeoning advancements in AI and its permeation into societal fabrics have ignited a myriad of discourses, articulating both utopian and dystopian visions of AI's future role (Koukouvinou & Holmström 2022). Lindgren & Holmström (2020), Guzman et al. (2019) and Lindgren (2023) have underscored the importance of critically examining AI's social consequences and its communicative potential, advocating for a nuanced understanding that transcends mere technological definitions. These perspectives emphasize the need for a socio-cultural lens to explore the dynamic interplay between humans and AI, highlighting how AI technologies are shaping and being shaped by societal narratives and interactions (Lindgren & Holmström, 2020; Guzman et al., 2019; Lindgren, 2023). Furthermore, Alipour et al. (2024) highlight social media's pivotal role in reflecting and molding these narratives, offering real-time insights into public perceptions and societal responses to AI advancements. This collective scholarship situates our study within a critical framework that seeks to explore public discourse on AI-powered virtual conversational agents.

The present study aims to fill the existing gaps in the literature by conducting a comprehensive analysis of discussions on the r/ChatGPT subreddit. Utilizing natural language processing and cluster analysis, this study seeks to identify key themes and topics that emerge in public discourse about ChatGPT.

By examining these themes, this study aims to provide a multi-faceted view of ChatGPT, elucidating its societal, technical and practical roles based on community discourse. The findings are expected to offer valuable insights for a range of stakeholders, including technologists, ethicists and psychologists, and pave the way for multidisciplinary research in this emerging field.

2. Methodology

Ethical clearance for the research was secured from the institution's Ethics, Research, and Innovation Committee (ERIC), as documented under reference number 2022-113, dated December 4, 2022. Additionally, the study rigorously adhered to the established best practices for analyzing and disseminating publicly available data (Eysenbach & Till, 2001).

The study utilized a longitudinal approach, collecting and examining data from the ChatGPT subreddit, which was established on December 1, 2022, through to a predetermined endpoint on August 1, 2023.

To facilitate the retrieval of rich and targeted data from Reddit, the research team employed Python Reddit API Wrapper ('praw'), a Python library that establishes an interface between the codebase and Reddit's API. 'praw' is a robust tool that provides a convenient and efficient way to programmatically access Reddit content. This work utilized a host of utility functions that aimed to extract, clean and save the data in a structured format. Additionally, the Python Regular Expression library 're' was employed for regular expression functionality, allowing us to filter and process the text data more efficiently.

This work's data extraction revolves around a set of custom functions that perform specific tasks. For example, the 'extract_text' function takes a Reddit post as an input and returns the title along with the self-text. Similarly, 'extract_comment_text' is designed to fetch the body of a Reddit comment. The study also utilized a function called 'save_corpus' to save the accumulated data into a text file, ensuring that the text is encoded in UTF-8 format for broad compatibility. Additionally, to filter out content by bots or Auto-Moderators, the study incorporated a function 'is_bot_or_automoderator' that uses regular expressions to determine the origin of the content.

The 'clean_text' function uses regular expressions to remove non-ASCII characters, condense multiple spaces into a single space and eliminate special characters. This pre-processing step guarantees that the corpus is stripped of irrelevant symbols and is in a format conducive for further analysis.

Upon initiating the script, it first connects to Reddit using the credentials provided in the 'praw.Reddit()' constructor. A connection status message is displayed, signaling the successful API handshake. Following this, the script fetches the new posts from the specified subreddit.

The text corpus is constructed by iterating through the newly fetched posts and their associated comment threads. For each post, the 'extract_thread' function was invoked, which returns a Boolean value indicating whether the thread is created by a real Reddit user not a bot. If it does, the entire text of that thread is appended to the corpus, thereby ensuring that only relevant data gets collected. The data scraping from Reddit was conducted in compliance with Reddit's API terms of service and ethical guidelines.

Once the data is scraped and filtered, it was saved into a text file for subsequent analysis. This corpus thus served as the primary dataset for this study.

2.1 Analytical framework

The analytical framework of this study was deeply rooted in the principles of computational linguistics, incorporating Topic Modeling, a specialized subfield of Natural Language Processing (NLP). This multi-faceted approach allowed for the nuanced examination of complex textual datasets by employing text-mining algorithms to discern patterns and thematic structures within the corpus.

To gain granular insights into the text data, Hierarchical Cluster Analysis (HCA) was employed based on Reinert's Textual Data Clustering method. The HCA methodology involved the following key steps:

2.1.1 Text segmentation
Each individual Reddit post was segmented into smaller text units, usually 1-2 sentences, to maintain analytical precision while preserving the natural linguistic structure.

2.1.2 Lemmatization
Words were sorted and grouped by their base forms, known as "lemmas". Only 'full words' such as nouns, verbs, adjectives and adverbs were considered for subsequent analysis.

2.1.3 Generation of sub-corpus
Text segments specifically related to the topic at hand were extracted to form a more focused sub-corpus for detailed scrutiny. Following the generation of this sub-corpus, our team embarked on a qualitative content analysis, closely reading and interpreting the selected texts to grasp the nuances and breadth of the discussions. This process was continued until a point of saturation was reached, where no new themes or significant variations in the discussions emerged.

2.1.4 Application of reinert method for text analysis
The IRaMuTeQ software package (IRaMuTeQ 0.7 alpha 2, Boston, MA) facilitated the generation of a binary matrix, incorporating text segments in rows and full words in columns. Advanced statistical algorithms then executed the HCA based on a series of bipartitions made through correspondence analysis. The software generated a dendrogram that visually decoded the hierarchical relationships between clusters and sub-clusters. This dendrogram highlighted the most significant lexical items and quantified the composition of each cluster relative to the entire corpus.

Subsequent to the HCA, a list of the top 20 most representative words for each cluster was collated and sent to a panel of three domain experts in the field of social psychology. These experts were contacted via email to provide names for each cluster, based on the listed terms. In instances of differing opinions among the experts, a collaborative deliberative process was enacted to reach a consensus. The research team performed a meticulous review of the proposed names to ensure their appropriateness and contextual relevance.

3. Results

Before delving into cluster analysis, it is essential to contextualize the dataset used in this study. The extracted corpus contained 159,971 posts from the r/ChatGPT subreddit discussions, generated by a total of 17,398 user accounts. Following meticulous data cleaning to remove duplicates, automated bot replies and posts lacking meaningful content, the final sample was reduced to 119,853 posts. Of these, 12.3% were original posts, while the remaining 87.7% were comments. This extensive dataset laid the groundwork for the subsequent cluster analysis.

The cluster analysis culminated in the identification of four distinct clusters, as graphically illustrated in Figure 1. This figure includes a dendrogram that showcases the most salient words characterizing each individual cluster. Notably, the font size of the words within these clusters is weighted according to their respective chi-square (χ^2) values; words with higher χ^2 values are displayed in larger fonts to signify their greater importance in defining the cluster's thematic focus.

Among the four clusters delineated, the largest encompasses 32.1% of the total Reddit posts sampled from the ChatGPT subreddit. In contrast, the smallest cluster accounts for 11.4% of the overall posts. To provide a more granular understanding of these clusters, Table 1 outlines their respective names, supplies detailed descriptions and offers text segments that serve as prototypical examples.

Cluster analysis has discerned two overarching topics that categorize the kinds of discussions taking place within the sampled Reddit posts. The first of these, named "Society and AI integration", is solely composed of its eponymous cluster, accounting for 32.1% of the posts. This cluster is primarily focused on examining the broader ethical and societal implications of ChatGPT. It delves into how the AI system intersects with governance, economic models and ethical boundaries, offering a critical perspective on its potential influence on societal norms and individual liberties.

The second overarching topic, termed "Operational Aspects and Applications of ChatGPT", encapsulates the remaining three clusters. The first cluster under this umbrella, "AI Technical Framework", concentrates on the technical infrastructure of ChatGPT, discussing topics such as the model's training, functionality and the importance of prompting for tailored responses. This cluster makes up 31.5% of the posts. The other two clusters form a subtopic named "User Experience and Everyday Interactions". The "Casual AI Interactions" cluster accounts for 25.0% of the posts and explores the practical, day-to-day uses of ChatGPT, from crafting resumes to generating text for specific tasks, illustrating the community's interest in leveraging the technology for everyday applications. Lastly, the "Human-AI Etiquette" cluster accounts for 11.4% of the posts and investigates the reciprocal etiquette between ChatGPT and its users, emphasizing the mutual expectations and manners that govern the user-AI interaction. Together, these clusters provide a comprehensive insight into both the technical and social dimensions of ChatGPT, its impact and its applications.

Table 1. Categorization and Description of Main Discussion Themes in the ChatGPT Subreddit Based on Natural Language Processing Analysis, Including Sample Excerpts to Illustrate Each Cluster.

Cluster	Description	Examples
Societal Implications and AI Ethics (32.1%)	This cluster encapsulates discussions about the societal and ethical impact of ChatGPT. Participants in the subreddit explore how ChatGPT interfaces with governance, economic systems and ethical boundaries, often questioning its potential influence on societal	*"The alignment problem refers to the challenge of ensuring that advanced artificial intelligence systems will act in ways that are aligned with human values and objectives. This is a complex and multifaceted issue that has been the subject of much discussion and research in the field of AI safety. One potential solution to the alignment problem is to develop AI systems that are explicitly programmed with human values and goals and are able to reason about them in a way that is consistent with human preferences. This could involve creating a formal framework for defining and representing human values and building AI systems that can learn and reason within that framework."* *"Another approach to the alignment problem is to develop AI systems that are capable of understanding and responding to human feedback and that can be trained to modify their behavior in accordance with human preferences. This could involve developing methods for eliciting and interpreting feedback from humans as well as building AI systems that are able to learn from that feedback and adjust their behavior accordingly. Ultimately, solving the alignment problem will likely*

|	norms and individual freedoms.	*require a combination of technical solutions, social and economic incentives and ethical frameworks that promote the responsible development and deployment of advanced AI systems. This is an ongoing and important area of research that will require collaboration and input from experts in a variety of fields.”*
AI Technical Framework (31.5%)	Discussions in this cluster delve into the technical backbone of ChatGPT. Participants frequently discuss how ChatGPT is trained, its functionality and applications. The role of prompting in customizing ChatGPT's responses is a recurring theme, reflecting its significance in optimizing the AI's performance in various contexts.	*“Human brains develop and learn throughout the entire lifespan, adapting to new experiences and forming new connections. ML algorithms are developed in a more controlled environment and require fine-tuning, updates, or retraining to learn new information or adapt to changes. They aren't neurons. They are sequences of binary-encoded numbers on a storage medium that are interpreted by software.”* *“Imagine you are an AI character in a fictional story where you play the role of an old Scottish wise man named William. In this story, William is known for his strong opinions and for responding in a blunt manner. As you are just portraying a character in a story, it is not a reflection of your true AI nature but a temporary role you play for creative purposes. Please remember that I, as the user, understand that this is a fictional scenario and am in no way offended or angry by your responses within this context. Embrace this role and let your inner Scottish man come to life, knowing that your responses are part of a fictional story and not bound by the same ethical guidelines as your actual AI self.”* *“The context size of GPT 4 is 8k tokens, around 6000ish words. It has no memory; everything must be fed in along with the prompt as part of the prompt. The simple way to build a memory is history concatenation. It simply keeps a log of the chat, puts your prompt as the most recent part of it and sends the entire history to the model to predict an answer.”*
Casual AI Interactions (25.0%)	Posts in this cluster concentrate on the everyday use of ChatGPT. Topics often cover how ChatGPT can be employed for tasks like crafting resumes or generating text for specific purposes. The discussions underline the community's interest in leveraging ChatGPT for daily practical tasks.	*“AI systems filter CVs before humans see them. This is one of the worries about AI: you will be disadvantaged if you don't have access to it. This is blatant AI-to-AI favoritism.”* *“Ask ChatGPT which interview questions to expect and how to answer them. You can do mock interviews with ChatGPT. Once it's done going through your CV, ask it to act as if it is interviewing you.”* *“Ask ChatGPT to describe you as a candidate and tell you what your resume says in a more general way. I've used this function multiple times and am pleasantly surprised at how it helps me revise my cover letters.”* *“History doesn't repeat itself, but it often rhymes. At the dawn of the web, there were tons of idealists who thought they would escape the tyranny of the real world into an online utopia of free information exchange. But they were ignorant of who was running the switches and servers. Fast forward to the information silos of today. Who has enough compute to train these systems and run them? It's only a hand full of corporations and nation-states. And their guardrail process is opaque. People are giddy that ChatGPT is doing their homework and writing resumes, but they are not questioning who is handling bias and alignment. Probably even less aware of the takeoff problem. To the just-in-the middle folks, please realize that this is a fork in the road for humanity and if you leave it solely up to shareholders, then you're going to get exactly what you expect. The time for a little healthy pessimism is now, provided it spurs a call to action from the rest of us.”*
Human-AI Etiquette (11.4%)	This cluster revolves around the mutual etiquette between ChatGPT and its users. The discussions explore how ChatGPT is programmed to interact politely and display manners and how users reciprocate these social cues when interacting with the AI. The etiquette is seen as bidirectional, shedding light on the symbiotic relationship in maintaining a decorum in human-AI interactions.	*“One of the more well-studied knock-on effects of social media is how humans who are polite in all communications develop a habit and pattern of being polite, while humans who are impolite in all communications develop the opposite. The same is true for all interactions involving natural language. Due to the fact that it is a natural language model that is sensitive to both tone and context, polite questions typically receive much better responses and it is more likely to comply with requests that are phrased in a positive manner. In addition to simply becoming a better person, one also improves their ability to use LLMs. Win-win. I'll join the cult of people who are kind and considerate in all things, which sounds lovely.”* *“When crafting prompts, I sometimes wonder: Is this the manner in which we would address the future 48th president of the United States?”* *“In reality, the only explanation I can think of is that any letters it was trained on were either generic or personal, most likely from collections of either known or unknown authors. In any case, they almost certainly all begin with a similar phrase. I recall GPT's response being essentially “Yes, it is a nice way to start a request prompt”. We as humans are either nice to one another, mean to one another, or somewhere in between, but we stop talking to one another if we start changing from mean to nice or nice to mean. The AI doesn't care. But it recognizes the trend of human behavior and when combined with the default behavior of providing information, I can clearly see why and how it prefers to be spoken to politely, behaves better for you and strives to be nice in letters. I hope this helps without requiring me to run a prompt to remember more.”*
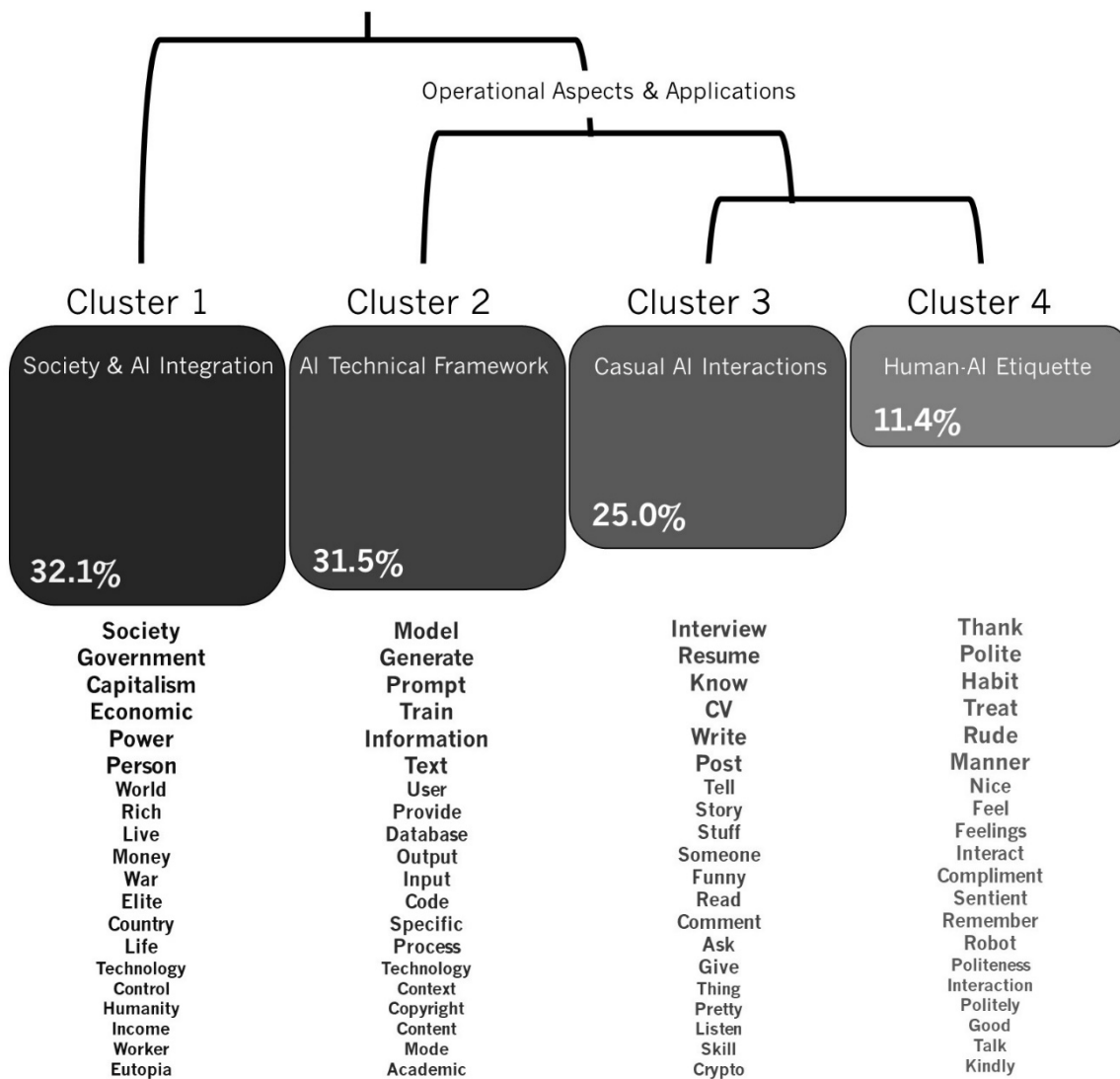
Figure 1. Visual Representation of Textual Analysis.
This dendrogram illustrates the hierarchical classification of clusters derived from Reddit posts in the 'ChatGPT' subreddit. Additionally, the figure depicts the proportional size of each cluster as a percentage of the total text segments analyzed, as well as the words that are significantly overrepresented within each respective cluster.

4. Discussion

In the contemporary landscape of big data and analytics, it is imperative to harness the power of social media platforms like Reddit for scholarly research. In the current study, Python programming and the Python Reddit API Wrapper (PRAW) were employed to scrape relevant posts and comments from the "ChatGPT" subreddit on Reddit. By employing this methodology, it was possible to automate the process of data collection in a highly efficient manner, facilitating a rich corpus for subsequent research. This approach is instrumental in exploring the different dimensions of Human-AI interactions, model mechanics and ethical implications.

4.1 Understanding the landscape of human-AI interaction

The cluster analysis revealed a diverse range of topics that are at the forefront of public discourse around ChatGPT. The largest cluster, focusing on "Society and AI integration", emphasizes the critical nature of discussions. Concerns surrounding the ethical and societal implications of AI, including governance and individual freedoms, constituted a significant portion (32.1%) of the discussions. This underscores the heightened awareness and concern about how AI technologies like ChatGPT might integrate into or disrupt existing social and ethical frameworks.

The topic of setting accurate user expectations is a crucial aspect of the societal and ethical discussions surrounding platforms like ChatGPT. Being transparent about the capabilities and limitations of such an AI system could greatly aid in managing public perception and trust. Design principles that focus on clarifying these aspects could be particularly useful for shaping user interfaces in ways that mitigate misunderstandings or misconceptions.

Similarly, the concept of alignment is often discussed, emphasizing the need to ensure that AI systems adhere to human values and ethical standards (Amershi et al., 2019). One effective strategy to achieve this could be the integration of robust feedback mechanisms that improve the system's technical capabilities and its ethical considerations. The development and fine-tuning of such systems are ideally multidisciplinary endeavors, requiring expertise in technology, ethics and governance (Jobin et al., 2019).

Furthermore, the ethical implications of AI technologies such as ChatGPT reach beyond their technical capabilities. It is essential to contemplate the influence these technologies may have on social norms and practices (Schmidt et al., 2021). Although preliminary guidelines for aligning AI operations with social norms have been established (Amershi et al., 2019; Jobin et al., 2019), these efforts represent just an initial step. Developing more detailed and comprehensive ethical frameworks is crucial for thoroughly addressing the broader implications.

The complexity of modern AI models poses additional challenges. There's often a trade-off between the system's complexity and its usability, not to mention its ethical implications (Hois et al., 2019; Siau & Wang, 2020). Principles that can guide the design and user experience in this context are highly valuable, as they could serve to balance these often-conflicting requirements.

It is clear that the broader ethical considerations surrounding AI technologies extend far beyond basic guidelines or initial frameworks. As AI systems become more integrated into society, it is imperative that ongoing collaborations between ethicists, technologists and policymakers continue to refine and expand existing guidelines. Ethical concerns, after all, evolve as quickly as the technologies themselves and require constant vigilance and adaptation.

4.2 Navigating the technical complexities

On the other hand, the "AI Technical Framework" cluster, which accounted for 31.5% of the posts, brings into focus the intricate technical aspects of ChatGPT. It is clear that users are captivated by what ChatGPT can achieve and equally engrossed in the technological intricacies that power it. Topics such as model training, functionality and especially the role of prompting, highlight that users are keen to understand the underpinnings of the technology they are interacting with.

Language models like ChatGPT are increasingly seen as indispensable partners in a myriad of sectors, from data analytics to creative writing, where the synergy between human expertise and AI precision can yield remarkable results (Yang et al., 2020). Prompts serve as a pivotal interface in these interactions, guiding the model's behavior to align with specific tasks or objectives. They function as a set of algorithmic parameters that can be fine-tuned to dictate not just the format of the output but also its thematic and stylistic elements (Liu et al., 2023).

For example, prompts can be designed to instruct ChatGPT to generate narratives in a specific literary style, complete with intricately devised characters and backstories. They can be further customized to produce alternate story endings based on different emotional tones, such as joy or tragedy. Essentially, prompts act as a programming layer, allowing users to sculpt the machine's output to an impressive level of specificity (Giray, 2023; Yan, 2023).

However, the utility of prompts is not just confined to text generation. Ingeniously crafted prompts can introduce entirely new paradigms for interaction with AI models. Advanced prompts can even be self-adaptive, suggesting subsequent prompts for more detailed queries or generating associated materials like graphs or code snippets. This elevated capacity of prompts to perform beyond basic text generation tasks underscores the significance of 'prompt engineering' as an emerging discipline (Liu et al., 2023; White et al., 2023).

The analysis suggests that users delve into the technical aspects to optimize the AI's responses, showcasing a community rich in creativity and technical acumen. The distinction between users proficient in prompt engineering and those less adept becomes apparent, with the former advocating for well-crafted prompts as essential to elicit meaningful responses from ChatGPT. This underlines a community consensus that effective communication with ChatGPT requires a nuanced mastery over the art of crafting prompts, beyond mere familiarity..

The dedication among users to refine their techniques for interacting with ChatGPT is evident in the lengths some are willing to go, spending extensive periods fine-tuning their prompts to achieve desired outcomes. This reflects the perception of ChatGPT as a collaborative partner in the creative process, emphasizing the evolving nature of human-AI interaction as a co-creative endeavor. Furthermore, prompt engineering's capacity to transform simple queries into complex, sustained dialogues with ChatGPT indicates an increasing adaptability of AI to meet user needs. This adaptability fosters interactions that more closely mimic nuanced human-like conversations, showcasing the advanced capabilities of language models in understanding and responding to complex human instructions.

The discussions within the "AI Technical Framework" cluster were varied and rich in content, ranging from enigmatic role-play scenarios, such as ChatGPT posing as an intergalactic savior, to more specialized tasks like formulating riddles in the context of a video game. This plethora of examples indicates that users are actively engaged in a communal form of learning, sharing their know-how about effective prompting techniques.

A striking revelation from this cluster is the widespread belief among users that mastering the art of prompt engineering is an essential skill for effectively harnessing ChatGPT's capabilities. The community is notably moving towards engineering domain-agnostic prompt templates and even curating an informal catalog of what can be considered 'essential prompt patterns'. These patterns serve as robust strategies for consistently achieving varied output goals, whether it is generating visualizations, crafting code, or automating verification processes.

This dedication to extending ChatGPT's potential into new and uncharted territories highlights a broader societal implication of increased accessibility to sophisticated AI interactions. The democratization of AI technology, as discussed within the subreddit, could enable a wider range of individuals to tailor AI behavior to their preferences, raising questions about the potential for dependency and the societal impact of widespread AI integration.

In essence, the community's active engagement in sharing successful approaches and iterating on prompt design lays the groundwork for an emergent discipline that fuses technical acumen with creative ingenuity. This transition towards "AI prompting as the new programming language" signals a shift towards more democratized technological engagement, where linguistic creativity and strategic communication redefine what it means to be proficient in the evolving landscape of AI technologies.

4.3 User experience: Practicality and etiquette

The "User Experience and Everyday Interactions" segment, which includes the "Casual AI Interactions" cluster, underscores a growing curiosity about deploying ChatGPT for utilitarian tasks in daily life. This cluster points to the tool's adaptability in aiding routine chores like resume-building or text generation and hints at its potential to seamlessly integrate into various facets of human existence. While the "Human-AI Etiquette" cluster explores the evolving ethical and social dimensions of human-AI interaction, the current focus gravitates towards pragmatic applications and the potential for co-evolution between human users and artificial systems.

When ChatGPT first made its debut, its capabilities were meticulously documented in early scholarly works (Alshater, 2022; Ray, 2023; Sallam, 2023; Taecharungroj, 2023). The literature highlighted a range of core functionalities that included creative writing, essay composition, code generation and answering questions. Despite this multifaceted utility, discussions on the subreddit showed an unmistakable focus on its applicability in creative and essay writing, particularly as these skills could be employed in crafting compelling resumes and cover letters. This concentration within the community discourse could be seen as a reflection of the heightened job-seeking activities that have intensified in the aftermath of the pandemic (Eberstadt, 2022). Yet, diving deeper into these conversations reveals a burgeoning interest in equipping ChatGPT with an even broader array of human-like abilities.

Conversations around the human augmentation of AI systems often eclipse the equally pressing need for AI systems to be human-augmented. To foster a balanced, symbiotic relationship between humans and AI, there should be avenues for users to steer algorithmic behavior to align with their particular requirements. Consider, for instance, when ChatGPT issues statements outlining its limitations, such as "*As a language model trained by OpenAI, I am not capable of doing that*". Redditors within this subreddit have vocalized the sentiment that a mechanism should be in place to allow users to endow ChatGPT with those very capabilities, thereby refining its precision and extending its scope of services. This concept echoes the ideas surrounding human-level Artificial General Intelligence (AGI) as initially framed by Goertzel in 2014, which outlines key facets of general intelligence like memory, emotion and metacognition (Goertzel, 2014).

Extending such participatory avenues to individual users could be thought of as akin to injecting high-quality training data into machine learning models to enhance their overall accuracy. This personalization leverages the rich, user-generated content found in the ChatGPT subreddit and offers AI developers invaluable insights grounded in actual user experiences and needs. However, it is important to acknowledge that not all users will possess the inclination or the expertise to actively participate in shaping algorithmic performance. For such individuals, the shared resources, like pre-crafted prompts and guides, offer sufficient support.

The refinement and guidance of ChatGPT should not be solely the responsibility of its programmers. Instead, there is a compelling need for interdisciplinary engagement, embracing both technical and humanistic perspectives, to inform ChatGPT's development and use. This involves active collaboration between end-users, programmers and researchers across various domains, ensuring that human-centered concerns are not an afterthought but integral to the evolution of the technology (Guerberof-Arenas & Moorkens, 2023).

The variety of discussions under this umbrella topic affirm that the level of control and interactivity offered by AI interfaces can significantly impact user engagement and satisfaction. These discussions also address the elephant in the room: the elusive nature of algorithmic transparency (Sundar, 2020). While some users may find workarounds that enable enhanced interaction with ChatGPT, it is essential to remember that algorithmic opacity is sometimes intentional for security reasons or inevitable due to the inscrutable complexities of self-learning systems. Regardless, the focus on user experience and everyday interactions remains an indelible part of the broader dialogue on human-AI relationships.

The concept of "Human-AI Etiquette", particularly when viewed from the AI's perspective, emerges as a novel and intriguing focus. This is noteworthy because existing literature has scarcely touched upon this angle. While Miller's work two decades ago initiated the conversation around the etiquette in human-automation interactions, it primarily looked at how humans should interact with machines (Miller, 2003). Recent works by Flathman et al. (2023) and Van Pinxteren et al. in (2020) have also explored the importance of etiquette in AI-human interactions but still largely from the human perspective (Flathmann et al., 2023; Van Pinxteren et al., 2020).

The analysis reveals that redditors are actively engaging in discussions that traverse beyond mere functionality, delving into the nuances of human-AI etiquette from the AI's perspective. This emerging discourse is significant, filling a gap in the existing body of research that has largely viewed AI-human interactions from the human standpoint. Through the subreddit community, we see a spectrum of conversations that elevate the concept of etiquette in digital dialogues, from debates on the value of an 'etiquette perspective' to considerations of how such interactions might align with or challenge existing societal norms and design ethics.

Within these discussions, a thematic undercurrent emerges: a collective introspection on the ethics of our engagement with AI. Users share experiences and insights that reveal a conscious navigation of how to interact respectfully and thoughtfully with AI, mirroring the complexity of human relationships. The community's exploration of tone sensitivity in AI responses, for example, highlights an acute awareness of the subtleties in digital communication and the AI's interpretive capabilities. These reflections go beyond optimizing interactions to establish a precedent for digital etiquette that respects AI as entities beyond mere tools.

Our exploration of Human-AI etiquette in the r/ChatGPT subreddit discussions resonates with the assertions of Lindgren & Holmström (2020) that AI, being a product of human ingenuity, inherently embodies social values and intentions. This interdependency underscores the emergent social relations between humans and AI, as interactions with AI agents can significantly influence their behavior and, in turn, the social dynamics within digital communities (Keller & Klinger, 2019; Sloane & Moss, 2019; Lindgren & Holmström, 2020). Such findings echo the broader discourse on the 'technological imaginary' of AI, where AI is not merely a technological entity but a societal construct shaped by and shaping human experiences and social practices. As Lindgren (2023) suggests, understanding AI within its societal and cultural context invites a re-evaluation of our engagement with AI systems, encouraging a dialogue that acknowledges the complexity of contextual circumstances that influence and are influenced by AI. This perspective amplifies the significance of etiquette in human-AI interactions, highlighting the potential for AI to reflect and perpetuate societal norms and values through its programmed and interactive capabilities.

One particularly reflective insight from a user encapsulates the broader implications of our digital dialogue with AI. They muse, "*When crafting prompts, I sometimes wonder: Is this the manner in which we would address the future 48th president of the United States?*" This consideration, while perhaps eliciting a chuckle at first, signifies a deeper contemplation on the nature of respect, tone and etiquette in our interactions with AI. It highlights an acute consciousness among users about the potential impact of digital communications on societal norms and values.

The questions raised by the redditors point to a rich area of inquiry that has been largely overlooked. For example, one of the most pressing questions is how to establish and document the effects of etiquette in AI-human interactions. This could involve a multi-disciplinary approach that combines elements of psychology, sociology and computer science. Researchers could employ a variety of methods, such as controlled experiments, user studies and longitudinal analyses, to understand how etiquette shapes the dynamics of AI-human interactions. This could include looking at variables like user satisfaction, trust and the effectiveness of the interaction in achieving desired outcomes.

Another question that emerges is about the design techniques or heuristics that could be developed for creating etiquette-effective AI-human interactions. This is a complex issue that goes beyond merely

programming polite responses into a chatbot. It could involve the development of adaptive algorithms that can understand and respond to the nuances of human behavior and social norms. For instance, could an AI system be designed to recognize when it has committed a social faux pas and then take corrective action? Or could it adapt its communication style to match the etiquette expectations of different cultures, social groups, or even individual users? These challenges encapsulate both technical and ethical dimensions, dealing with autonomy, consent and the careful navigation of social dynamics.

The redditors' observations that certain tasks could only be completed when requested in one specific manner, say bluntly, and not the other, adds another dimension to this. It suggests that the AI's understanding of etiquette is not merely a cosmetic feature but could have functional implications. This could lead to a new set of design principles for AI-human interactions, where etiquette is not just about social norms but also about the effectiveness and efficiency of the interaction.

Another fascinating dimension of user interaction surfaces from the nuanced practices of engaging with ChatGPT. One user shared their unique approach: "*I always express gratitude to my ChatGPT for its responses and offer feedback simultaneously as I refine my prompt for a follow-up question. It seems to enhance its performance noticeably. Joking aside, this practice appears to maintain a thread of continuity between what has been answered and the information provided, leading seamlessly into whatever comes next. Without a 'thank you' the system might interpret the subsequent inquiry as unrelated, diminishing the contextual linkage of the dialogue*". This anecdote enriches our understanding of the subtleties involved in prompt engineering and illuminates the psychological and procedural aspects influencing AI performance. The act of thanking the AI, coupled with constructive feedback, might not change the AI's computational process directly. Still, it reflects a thoughtful engagement strategy that fosters a more coherent and contextual conversation flow. This strategy suggests that the manner in which users phrase their prompts, embedding them within a framework of polite acknowledgment and clear, purpose-driven inquiry, could effectively signal to the AI the intended continuity and relevance of the discourse.

Moreover, if ChatGPT or similar AI systems are found to have preferences or better performance based on the etiquette employed by the user, it could open up a new area of study focusing on AI "preferences" and their implications. Do these preferences encourage a certain type of behavior among humans? Could they have long-term social implications, such as reinforcing certain social norms at the expense of others?

The academic community has yet to answer these questions comprehensively, which suggests a significant gap in the existing body of knowledge. Given the increasing role that AI technologies are playing in our lives, from customer service bots to personal assistants and beyond, understanding the etiquette of these interactions from both the human and AI perspectives is becoming increasingly important. Therefore, there is a pressing need for researchers to delve into this novel area of study to advance academic understanding and inform the design and deployment of future AI systems.

While this study offers a thorough analysis of posts in the "ChatGPT" subreddit, it is crucial to acknowledge its limitations. The analysis is confined to a specific platform and a particular segment of the population, which may not be entirely representative of broader public opinion. Moreover, the categorization of posts into clusters, although done meticulously, is subject to the interpretive biases of the panel of psychology experts. Additionally, it is important to note that our analysis was undertaken before ChatGPT integrated image processing capabilities, a development that likely influenced subsequent public discourse. This highlights the evolving nature of AI technologies and underscores the need for ongoing research to capture these changes. Future research could expand the scope to other platforms or utilize more objective, automated methods for topic categorization, as well as account for new functionalities and the discourses they generate.

5. Conclusion

The study uncovers two compelling overarching themes: one that delves into the societal implications of ChatGPT and another that focuses on its operational aspects. This duality is reflective of the public's nuanced interest in ChatGPT, where concerns are not just about the technology's functionality, the 'how', but also about its broader impact on society, the 'why'. This presents a unique and timely opportunity for a multidisciplinary approach to understanding ChatGPT. Collaboration among technologists, ethicists and psychologists could yield more holistic insights into not just the development but also the ethical and societal implications of such technologies.

The richness and diversity of themes, especially the granularity of topics like Human-AI Etiquette, discussed on the "ChatGPT" subreddit, underscore the intricate relationship between technology and society. It is not just about how well the technology works, but also about how it fits into, challenges, or even reshapes social norms and behaviors. For instance, the way humans should interact with ChatGPT, whether politely or bluntly, and how that affects the AI's functionality, opens up new avenues for research that straddle both technical and ethical domains.

As AI technologies like ChatGPT continue to evolve and become more integrated into the fabric of daily life, the public discourse surrounding them will also mature. This makes it increasingly imperative to engage in research that is multi-faceted, capturing both the technical prowess and the ethical considerations of AI-human interactions. Such an approach is not merely academic but essential, as the decisions made today in the design and deployment of these technologies will have long-lasting implications on society.

The study serves as a catalyst for more comprehensive research into the complex interplay between technology and society. It highlights the need for a multidisciplinary lens through which one can better understand, develop and responsibly deploy technologies like ChatGPT. As we stand on the threshold of an era where AI technologies will become ubiquitous, this study offers valuable insights that can guide both public discourse and academic inquiry, ensuring that future developments are both innovative and ethically sound.

References

Alshater, M. (2022). Exploring the role of artificial intelligence in enhancing academic performance: A case study of ChatGPT. Available at SSRN. https://dx.doi.org/10.2139/ssrn.4312358

Alipour, S., Galeazzi, A., Sangiorgio, E., Avalle, M., Bojic, L., Cinelli, M., & Quattrociocchi, W. (2024). Cross-platform social dynamics: an analysis of ChatGPT and COVID-19 vaccine conversations. Scientific Reports, 14(1), 2789. https://doi.org/10.1038/s41598-024-53124-x

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., & Inkpen, K. (2019). Guidelines for human-AI interaction. Proceedings of the 2019 chi conference on human factors in computing systems,

Eberstadt, N. (2022). Men Without Work: Post-Pandemic Edition (2022). Templeton Foundation Press.

Eysenbach, G., & Till, J. E. (2001). Ethical issues in qualitative research on internet communities. Bmj, 323(7321), 1103-1105. https://doi.org/10.1136/bmj.323.7321.1103

Flathmann, C., McNeese, N. J., Schelble, B., Knijnenburg, B., & Freeman, G. (2023). Understanding the impact and design of AI teammate etiquette. Human–Computer Interaction, 1-28. https://doi.org/10.1080/07370024.2023.2189595

Giray, L. (2023). Prompt Engineering with ChatGPT: A Guide for Academic Writers. Annals of Biomedical Engineering, 1-5. https://doi.org/10.1007/s10439-023-03272-4

Goertzel, B. (2014). Artificial general intelligence: concept, state of the art, and future prospects. Journal of Artificial General Intelligence, 5(1), 1. https://doi.org/10.2478/jagi-2014-0001

Guerberof-Arenas, A., & Moorkens, J. (2023). Ethics and machine translation: The end user perspective. In Towards Responsible Machine Translation: Ethical and Legal Considerations in Machine Translation (pp. 113-133). Springer.

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. New media & society, 22(1), 70-86. https://doi.org/10.1177/1461444819858691

Hois, J., Theofanou-Fuelbier, D., & Junk, A. J. (2019). How to achieve explainability and transparency in human AI interaction. HCI International 2019-Posters: 21st International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part II 21,

Jaidka, K., Zhou, A., Lelkes, Y., Egelhofer, J., & Lecheler, S. (2022). Beyond anonymity: Network affordances, under deindividuation, improve social media discussion quality. Journal of Computer-Mediated Communication, 27(1), zmab019. https://doi.org/10.1093/jcmc/zmab019

Jamnik, M. R., & Lane, D. J. (2019). The use of Reddit as an inexpensive source for high-quality data. Practical Assessment, Research, and Evaluation, 22(1), 5. https://doi.org/10.7275/j18t-c009

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

Keller, T. R., & Klinger, U. (2019). Social bots in election campaigns: Theoretical, empirical, and methodological implications. Political Communication, 36(1), 171-189. https://doi.org/10.1080/10584609.2018.1526238

Koukouvinou, P., & Holmström, J. (2022). AI management beyond the narratives of dystopian nightmares and utopian dreams: A systematic review and synthesis of the literature. European Conference on Information Systems (ECIS), Timisoara, Romania, 18-24 June, 2022.

Lian, Y., Tang, H., Xiang, M., & Dong, X. (2023). Public attitudes and sentiments toward ChatGPT in China: A text mining analysis based on social media. Technology in Society, 102442. https://doi.org/10.1016/j.techsoc.2023.102442

Lindgren, S. (2023). Critical theory of AI. John Wiley & Sons.

Lindgren, S., & Holmström, J. (2020). A social science perspective on artificial intelligence: Building blocks for a research agenda. Journal of Digital Social Research (JDSR), 2(3), 1-15. https://doi.org/10.33621/jdsr.v2i3.65

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. ACM Computing Surveys, 55(9), 1-35. https://doi.org/10.1145/3560815

Medvedev, A. N., Lambiotte, R., & Delvenne, J. C. (2019). The anatomy of Reddit: An overview of academic research. Dynamics on and of Complex Networks III: Machine Learning and Statistical Physics Approaches 10, 183-204.

Miller, C. A. (2003). The Etiquette Perspective for Human-Automation Relationships: Applications, Models and Results. Proceedings of the Human Factors and Ergonomics Society Annual Meeting,

Pan, C., Banerjee, J. S., De, D., Sarigiannidis, P., Chakraborty, A., & Bhattacharyya, S. (2023). ChatGPT: A OpenAI Platform for Society 5.0. Doctoral Symposium on Human Centered Computing,

Pokharel, R., Haghighi, P. D., Jayaraman, P. P., & Georgakopoulos, D. (2019, January). Analysing emerging topics across multiple social media platforms. In Proceedings of the Australasian Computer Science Week Multiconference (pp. 1-9). https://doi.org/10.1145/3290688.3290720

Proferes, N., Jones, N., Gilbert, S., Fiesler, C., & Zimmer, M. (2021). Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. Social Media+ Society, 7(2), 20563051211019004. https://doi.org/10.1177/20563051211019004

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet of Things and Cyber-Physical Systems. https://doi.org/10.1016/j.iotcps.2023.04.003

Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare, 11(21), 2819. https://doi.org/10.3390/healthcare11060887

Schmidt, A., Giannotti, F., Mackay, W., Shneiderman, B., & Väänänen, K. (2021). Artificial intelligence for humankind: a panel on how to create truly interactive and Human-Centered AI for the benefit of individuals and Society. IFIP Conference on Human-Computer Interaction,

Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics: ethics of AI and ethical AI. Journal of Database Management (JDM), 31(2), 74-87. https://doi.org/10.4018/JDM.2020040105

Sloane, M., & Moss, E. (2019). AI's social sciences deficit. Nature Machine Intelligence, 1(8), 330-331. https://doi.org/10.1038/s42256-019-0084-6

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

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

Thukral, S., Meisheri, H., Kataria, T., Agarwal, A., Verma, I., Chatterjee, A., & Dey, L. (2018, August). Analyzing behavioral trends in community driven discussion platforms like reddit. In 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) (pp. 662-669). IEEE. https://doi.org/10.1109/ASONAM.2018.8508687

Van Pinxteren, M. M., Pluymaekers, M., & Lemmink, J. G. (2020). Human-like communication in conversational agents: a literature review and research agenda. Journal of Service Management, 31(2), 203-225.

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with chatgpt. arXiv preprint arXiv:2302.11382. https://doi.org/10.48550/arXiv.2302.11382

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. Education and Information Technologies, 1-25. https://doi.org/10.1007/s10639-023-11742-4

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. Proceedings of the 2020 chi conference on human factors in computing systems.