Can (A)I Have a Word with You? A Taxonomy on the Design Dimensions of AI Prompts

Braun, Marvin; Greve, Maike; Kegel, Felix; Kolbe, Lutz M.; Beyer, Philipp Emanuel

Document Version
Final published version

Published in:
Proceedings of the 57th Annual Hawaii International Conference on System Sciences, HICSS 2024

DOI:
10125/106443

Publication date:
2024

Can (A)I Have a Word with You? A Taxonomy on the Design Dimensions of AI Prompts

Marvin Braun
University of Göttingen
marvin.braun@uni-goettingen.de

Maike Greve
University of Göttingen
maike.greve@uni-goettingen.de

Felix Kegel
University of Göttingen
felix.kegel@uni-goettingen.de

Lutz M. Kolbe
University of Göttingen
lutz.kolbe@wiwi.uni-goettingen.de

Philipp Emanuel Beyer
University of Göttingen
p.beyer@stud.uni-goettingen.de

Abstract

Emerging generative artificial intelligence (AI) is revolutionizing human collaboration with machines through their remarkable processing of human language. Consequently, the role of AI is transitioning as it becomes adaptable to specific tasks through human language, without the need for additional training data. These "prompts" enable end-users to fully harness the potential of AI. However, the process of designing prompts is currently unstructured, relying mostly on trial-and-error processes. Studies emphasize the significant impact of prompt design on the AI's generated results. Therefore, there is a clear need to establish a structured approach to prompt design. To address this need, we analyzed scientific literature and practical examples to develop a taxonomy of prompt design. The taxonomy serves as a valuable resource for both practitioners and researchers, offering an overview of current AI prompt approaches and shedding light on potential areas for future research.

Keywords: generative artificial intelligence, prompts, taxonomy.

1. Introduction

Recent advancements in the field of generative artificial intelligence, such as Large Language Models (LLMs) and Generative Pre-trained Transformer models (GPT), have significantly enhanced the capabilities of this technology (Bommasani et al., 2021). Today, generative AI applications such as ChatGPT and Midjourney can execute complex and contextual tasks such as code debugging (Zamfirescu-Pereira et al., 2023), creating images (Oppenlaender, 2022), music (Guo et al., 2023), videos (Yan et al., 2021), and texts (Eloundou et al., 2023). Thereby, tasks are executed within seconds and match or even surpass human abilities (Eloundou et al., 2023; Wu et al., 2022).

These advances are largely facilitated by the increased ability of AI to process and generate human language (Wu et al., 2022). This not only opens up new avenues for human collaboration with AI-based systems but also introduces the concept of *prompting* (Bommasani et al., 2021; Liu et al., 2021). Building upon the framework proposed by Bach et al. (2022), we define 'prompting' as the process of instructing a generative AI to perform a task using natural language. In this context, we define prompt design (also referred to as prompt engineering) as the activities of tailoring a prompt to a specific task and utilizing the possibilities of the respective generative AI model.

Thereby, prompting allows generative AI to adapt to specific tasks, eliminating the need for additional fine-tuning with specialized data (Reynolds & McDonell, 2021). Earlier AI models were specifically fine-tuned for fixed tasks, in contrast, a generative AI model can be adapted to these tasks through prompts, eliminating the need for additional training iterations (Bommasani et al., 2021). Consequently, unlike its predecessors, generative AI is capable of performing a wider range of tasks and is not confined to a specific, narrow domain (Eloundou et al., 2023).

However, designing prompts that maximize the benefits of generative AI remains a challenge (Dang et al., 2022). Given the absence of fixed patterns for prompt design and the ease of formulating them through natural language, designing prompts is largely a trial-and-error process including a high degree of arbitrariness (Zamfirescu-Pereira et al., 2023). Moreover, AI models will not hint at potential better-fitting prompts but just execute the requested task, keeping users in suspense whether their prompt could have been improved to achieve better results. To avoid these pitfalls and increase the quality of the generated results, the practice of prompt engineering emerged

URI: https://hdl.handle.net/10125/106443
978-0-9981331-7-1

(Zamfirescu-Pereira et al., 2023). Studies have recently begun to investigate single components of prompts and their impact on produced results (Liu et al., 2021). For instance, a recent work indicates that longer prompts can result in worse AI performance (Wu et al., 2022). Moreover, information included in prompts needs to be aligned with each other to increase the AI performance (i.e., vice versa, contradicting information decrease result quality) (Wu et al., 2022). However, we are still lacking a clear overview of possible design dimensions that can be addressed when formulating prompts. While the phenomenon is still growing in scientific relevance, there is a need to conceptualize the growing corpus of both scientific literature and practical examples to enable systematic research. An initial classification would be beneficial for researchers to further study the different aspects of prompts. We argue that this phenomenon is particularly suited for the information systems (IS) community to investigate further. First, IS research provides extant literature on human-AI interaction (e.g., Rzepka & Berger, 2018) that can be utilized to structure prompts. Second, prompts are of particular importance for future human-AI interaction as they open new interaction patterns and system design based on generative AI. Hence, the topic of prompt design is not only of practical relevance but has also led IS scholars to increasingly call for an investigation of the impact of prompt design (Dwivedi et al., 2023). Based on this motivation, in this paper, we address the following research question:

'Which design dimensions and characteristics should be considered when creating AI prompts, and how?'.

To answer this research question, we develop a taxonomy of AI prompts design dimensions, following the taxonomy development method proposed by Nickerson et al. (2013) and the revised method by Kundisch et al. (2022). Moreover, through an initial application of the taxonomy, we are able to demonstrate its usefulness and versatility.

The paper is structured as follows: First, we provide a theoretical background on human-AI interaction and prompts. Next, we describe our research design as well as the analysis of academic literature and individual objects. Consequently, we derive a taxonomy, apply it to three example prompts and discuss the design dimensions in light of theoretical contributions and practical implications. Lastly, we outline the limitations of this work and suggest a future research agenda for AI prompts based on our taxonomy.

2. Conceptual Background

2.1. A Brief History of Prompting

A vital component of AI is its capacity to process human language, a field known as natural language processing (NLP). NLP encompasses numerous applications, including dialogue systems embedded in conversational agents and digital assistants (Maedche et al., 2019), as well as text mining and sentiment analysis (Hirschberg & Manning, 2015). To accomplish this, NLP leverages concepts from machine learning, such as deep neural networks (Dellermann, Calma, et al., 2019). Traditionally, AI models have been contextualized through training on specific subsets of data for fixed purposes, a process often referred to as fine-tuning (Wu et al., 2022) (see Figure 1).

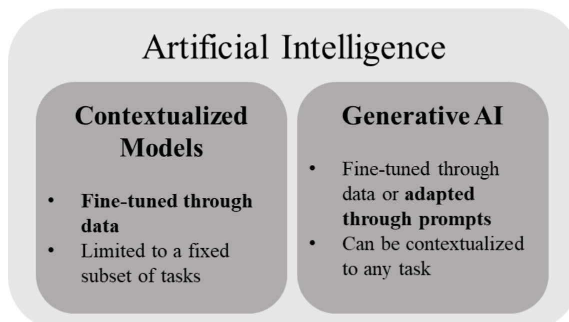

Figure 1. Contextualized Models and Generative AI. Adapted from Dellermann et al. (2019).

The emergence of generative AI, coupled with advancements like LLMs, introduced a paradigm shift. Generative AI is trained on considerably larger and more diverse corpora, encompassing text, images, speech, and signal data (Bommasani et al., 2021). As a result, such models offer multiple advantages over conventional, contextualized models. First, they are not limited to a particular domain but are general models that can be adapted to a wide range of contexts, such as finance or healthcare (Bommasani et al., 2021). Second, within these contexts, generative AI is not tied to a specific task but can handle a wider array of tasks without additional fine-tuning (Wu et al., 2022). This adaptability is achieved through natural language prompts (Wu et al., 2022), eliminating the necessity for extensive domain-specific training data (Brown et al., 2020). Previously, developers had to undertake these training processes (Bommasani et al., 2021). This allows for novel ways of interacting and working with AI-based IS through natural communication (Myers et al., 2000).

Table 1. Example prompts.

ID	Prompt
#1	Explain the concept of [XY] in simple terms.
#2	I want you to act as a spoken English teacher. I will speak to you in English, and you will reply to me in English. I will provide you with an example of how my text should be formulated. Go through every mistake in detail and reason why it is a mistake. I want you to ask me a question in your reply.
#3	portrait art of blade runner, 8 k ultra realistic, atmosphere, intricate, full of colour, cinematic lighting, trending on artstation, hyperrealistic, focused, extreme details, unreal engine 5

Since this work does not concentrate on technical aspects but rather on human-AI interaction, in the following, we categorize generative AI and its approaches, such as LLMs, under the term AI for simplicity.

2.2. Prompting as a Central Process of Human-AI Interaction

Following Terveen's (1995) definition, human-AI interaction can be conceptualized as "two or more agents [that] work together to achieve shared goals […] involving at least one human and at least one computational agent" (p. 67). This interaction is primarily influenced by the AI, the human user, and the surrounding context, which collectively shape the outcomes (Rzepka & Berger, 2018). Until recently, human-AI interaction was typically achieved through applications that incorporated AI into IS to carry out a fixed range of tasks in tandem with humans (Baird & Maruping, 2021).

With the ability to interact with AI through prompts, new interaction possibilities have become available; heralding a change of interaction where the AI becomes agentic in its behavior and transitions its role towards that of a digital assistant as projected by Baird and Maruping (2021) and Maedche et al. (2019). While there is extant literature on the transition from traditional human-computer-interaction to human-AI interaction and its implications for the research domain of IS (e.g., Rzepka & Berger, 2018; Seeber, Bittner, et al., 2020), we are currently lacking a structured approach to further analyze this new form of interaction (i.e., prompts) whose design is not a trivial task, as it might initially appear. Prompts offer unlimited design choices and AI will most of the time generate some result without hinting if there are potential improvements.

Works from the domain of computer science indicate large quality variations depending on the design of prompts (Dang et al., 2022; Sorensen et al., 2022) and have established the domain of prompt design (also referred to as prompt engineering) (Liu et al., 2021) investigating single aspects of prompts such as the quantity of provided examples (Dang et al., 2022), the length of prompts (Wu et al., 2022) and the usage of magic keywords (i.e., terms that are mostly used in text-to-image generation to achieve better image results) (Oppenlaender, 2022). Nonetheless, there are currently no standardized processes or frameworks on prompt design to serve as blueprints; hence, it is primarily a process of trial and error (Zamfirescu-Pereira et al., 2023). Finally, the design options to formulate prompts are almost infinite due to their reliance on human language (see Table 1). Therefore, we argue that it is crucial to understand for IS research how prompts are impacting AI performance and which possibilities are currently emerging from other research domains such as computer science to investigate this emerging phenomenon in more detail.

3. Research Design

3.1. Taxonomy Development Method

This study aims to shed light on relevant dimensions and characteristics of AI prompts that can be utilized by users and developers to adapt AI to their specific needs. To achieve this, we use the extended method of taxonomy development proposed by Kundisch et al. (2021), which builds on Nickerson et al. (2013). This process comprises several steps. Initially, we define the object of analysis (1), identify the target user groups (2), and specify the taxonomy's purpose (3). We examine prompts for publicly available AI models, which are generally used for interaction and collaboration with said models. Our target user group is twofold. We provide aim to provide a comprehensive overview for researchers but also users of said AI. The taxonomy's purpose is to provide a structured approach to prompts, thereby highlighting future research avenues for this burgeoning phenomenon and offering guidance to prompt engineers. Subsequently, we establish the meta-characteristic (4) of our taxonomy as 'design dimensions of prompts that shape the AI's resulting output'. Last, we employed the following ending conditions (5). As objective conditions we defined that all selected papers and objects need to be analyzed. Moreover, during the last iteration, no dimension or characteristic is added and every characteristic is

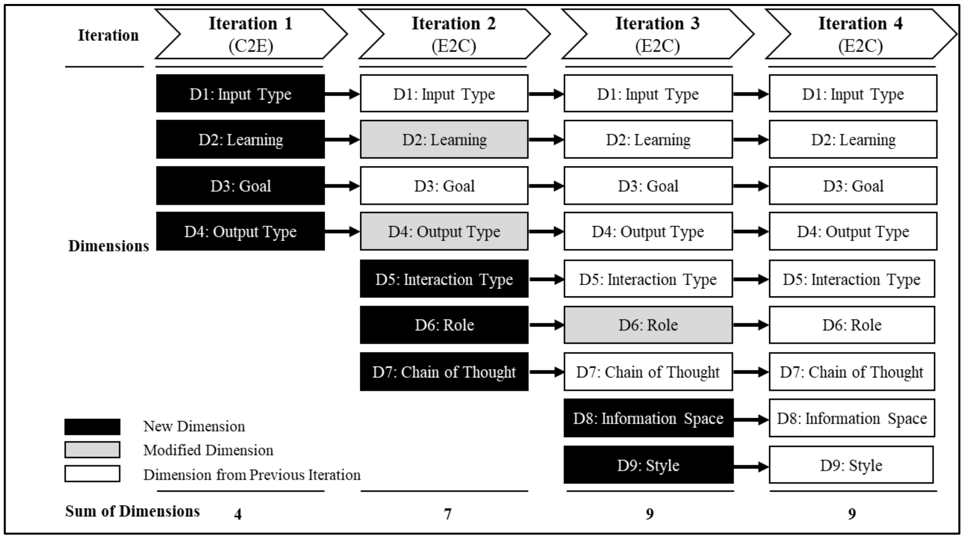

Figure 2. Taxonomy development process.

unique within its dimension. As subjective conditions we defined that the taxonomy needs to be concise (only relevant dimensions and characteristics are included), extendible (dimensions and characteristics can be added to the taxonomy), robust (objects can be differentiated by the taxonomy), comprehensive (all objects can be classified by the taxonomy) and explanatory (the taxonomy explains the object) (Nickerson et al., 2013).

Following the proposed method of Nickerson et al. (2013), we conducted four iterations to develop the taxonomy (see Figure 2). We started with a conceptual-to-empirical iteration to utilize profound knowledge of scientific literature as a starting point. We reviewed literature from the research domains of information systems and computer science to derive initial dimensions and characteristics. We identified prompts that were included in this scientific literature. As a second iteration, we selected an empirical-to-conceptual approach. Based on the database of prompthero.com, we classified the top 100 best-rated prompts and identified three new dimensions and characteristics. In the third iteration (again empirical-to-conceptual), we utilized github.com to identify more curated and highlighted prompts. In this iteration, 36 additional prompts were analyzed, and two dimensions were added. Lastly, we undertook a fourth iteration (empirical-to-conceptual), extracting the 40 most recent prompts from prompthero.com. After satisfying all ending conditions in the fourth iteration, we concluded the taxonomy development process.

3.2. Data Collection Process

Literature Review: We chose a conceptual-to-empirical cycle for our first iteration to develop a profound knowledge of AI prompts. We conducted a systematic literature review (Webster & Watson, 2002) to grasp all relevant papers from the different research streams and structure their insights After an initial unstructured search (Google Scholar and Web of Science), we selected multiple databases to ensure we include the most important papers. We specifically chose the AIS eLibrary, IEEE, Web of Science, and EBSCOhost. We considered journal articles as well as conference proceedings due to the novelty of the topic. Based on our unstructured search, we defined the following search string: “(prompt* OR fine-tun*) AND (AI OR artificial intelligence OR LLM OR Large Language Model OR Machine Learning OR ML)”. This search yielded 864 search results, which were screened for exclusion. After an initial screening of the abstract and titles, we conducted a full-text screening. Papers were specifically included if they dealt with the design of prompts or investigated single components of prompt design such as the impact of zero- or few-shot learning on the output quality. Moreover, we excluded paper that evaluated or benchmarked the overall AI performance as this is not relevant for the design of prompts. Thereby, we identified 28 papers that research prompt design.

Prompt Review: We considered prompts from multiple sources. During the first iteration, we included the prompts we found in the analyzed papers. Afterward, we included databases for prompts in the subsequent iterations (i.e., prompthero.com, promptdb.ai). We specifically chose these databases because they offer a systematic way for searching and filtering prompts. We aimed to create a dataset that embraces a wide spectrum of diversity. This involved encompassing all model types present in the respective databases (e.g., image generation, text generation, role play, etc.) to fortify the robustness of our systematic approach. If there were multiple highly similar prompts, we only included one prompt in our database. In total, we collected 176 prompts and stopped when the taxonomy was not further extended, however, it should be noted that there is a plethora of other sources that could be considered for prompt collections such as blogs and further prompt databases that are currently emerging.

To analyze the content of the prompts, we employed the content analysis method proposed by Mayring (2015). Given the emerging nature of prompt design and its limited conceptual foundation, we used inductive category development (i.e., dimensions and characteristics). To test our approach a subset of 20 prompts was selected first and analyzed with inductive category development by one author. The results were then discussed and validated by the author team to increase the rigor of this approach.

4. Taxonomy on Design Dimensions of AI Prompts

The resulting taxonomy comprises nine dimensions, grouped into three meta-dimensions (see Table 2). Drawing on established concepts of human-AI interaction (Rzepka & Berger, 2018), we derived the meta-dimensions of interaction, context, and outcome. We chose to include these meta-dimensions as they group the dimensions according to their impact on human-AI interaction. Moreover, these dimensions outline the generic options for designing prompts for AI models and are not constrained to a specific type of LLM, such as ChatGPT or Midjourney.

4.1. Interaction

The interaction meta-dimension includes the design dimensions that shape the communication process between the AI and the user. It encompasses the modes and methods through which information is exchanged, shaping the resulting interaction between humans and AI.

Input and Output Type: Through the ability of AI to process unstructured information, humans can interact with AI as they would with humans (Maedche et al., 2019; Seeber, Waizenegger, et al., 2020). Works on LLMs and prompt design emphasize the importance of preparing the AI to process heterogenic information and provide outputs defined through the prompts (Bommasani et al., 2021; Qiao et al., 2022). Moreover, they are able to transform the information input type to another output type (e.g., text-to-image - Oppenlaender, 2022, or text-to-code – Bommasani et al. 2021). We identified the input and output types of text, image, audio, video, and other. Both dimensions are non-exclusive because they can also be combined (e.g., a textual prompt that describes how to process a passed image).

Interaction Type: The interaction between humans and AI can be defined through the prompt. Following works on human-AI collaboration (Dellermann, Ebel, et al., 2019; Wiethof & Bittner, 2021) the interaction can be classified as human-in-the-loop or computer-in-the-loop. While there are further sub-patterns of collaboration, depending on the respective task and context, we will focus on these two

Table 2. Taxonomy on design dimensions of AI prompts.

<table>
<tr><th colspan="3">Dimensions</th><th colspan="30">Characteristics</th></tr>
<tr><td rowspan="3">Interaction</td><td>Input Type</td><td rowspan="2">NE</td><td rowspan="2" colspan="6">Text</td><td rowspan="2" colspan="6">Image</td><td rowspan="2" colspan="6">Audio</td><td rowspan="2" colspan="6">Video</td><td rowspan="2" colspan="6">Other</td></tr>
<tr><td>Output Type</td></tr>
<tr><td>Interaction Type</td><td>NE</td><td colspan="15">Computer-in-the-loop</td><td colspan="15">Human-in-the-loop</td></tr>
<tr><td rowspan="4">Context</td><td>Role</td><td rowspan="2">ME</td><td rowspan="2" colspan="15">Not defined</td><td rowspan="2" colspan="15">Defined</td></tr>
<tr><td>Style</td></tr>
<tr><td>Learning</td><td>ME</td><td colspan="10">Zero-shot</td><td colspan="10">One-shot</td><td colspan="10">Few-shot</td></tr>
<tr><td>Information Space</td><td>ME</td><td colspan="10">Not defined</td><td colspan="10">Explicit internal</td><td colspan="10">Explicit external</td></tr>
<tr><td rowspan="2">Outcome</td><td>Chain of Thoughts</td><td>ME</td><td colspan="15">Single step</td><td colspan="15">Step-by-step</td></tr>
<tr><td>Goal</td><td>NE</td><td colspan="5">Learn</td><td colspan="5">Lookup</td><td colspan="5">Investi-
gate</td><td colspan="5">Monitor /
extract</td><td colspan="5">Decide</td><td colspan="5">Create</td></tr>
<tr><td colspan="33">ME = Mutually Exclusive, NE = Non-Exclusive</td></tr>
</table>

general characteristics due to the generalizability of the taxonomy. During a computer-in-the-loop interaction, the AI acts as an assistant or tool to support and augment human decision-making or tasks (Lubars & Tan, 2019). It assists users by providing suggestions, recommendations, or performing tasks under human supervision. In contrast, during a human-in-the-loop interaction, the AI system operates more autonomously and provides output for the user, and requests information when appropriate (Lubars & Tan, 2019). Following this, prompts can be designed to either put the human or the computer in the loop by defining who conducts the main task and who gives additional input. For instance, a prompt that involves the interaction type of computer-in-the-loop is the following:" I will provide you with an e-mail and want you to give me feedback on the text and make suggestions on it". On the other hand, a prompt that includes human-in-the-loop interaction might be: 'Please write a script for a webinar on the best practices for using our product [XY]. I will provide iterative feedback, and you will adapt it based on the feedback,' emphasizing the distribution of tasks between humans and AI.

4.2. Context

This meta-dimension encompasses design mechanisms in which the AI's responses can be adjusted, fine-tuned, and adapted to a specific context. It includes the AI's role in the interaction, the thought process it simulates, the stylistic elements it incorporates, and the information space that can be defined by the user.

Role: The general behavior of the AI can be tailored to an individual context through the inclusion of a specific role. Because the possibilities of including personas are unlimited, we derived the general characteristics of defined and not defined. Interestingly, we found only one scientific paper that mentioned roles as a way of designing prompts (White et al., 2023). However, most prompts that we analyzed included roles. These roles typically describe jobs such as social media influencers, software developers, or psychologists (e.g., PromptHero, 2023). Following the work of White et al. (2023), including roles (also referred to as personas) enables users to change the AI's general behavior and output without specifying it in detail. This design element can be preferable if the exact desired outcome is unknown.

Style: In contrast to the role, which adapts the whole AI to a specific task, single modifiers (styles) can be utilized to refine the AI's output on a more granular level. Similar to the role, styles can be defined or not defined. Previous work shows that keywords (also referred to as magic terms) can be utilized for images, e.g., to specify the art style of an image as 'oil painting') (Oppenlaender, 2022). Moreover, empirical studies suggest that styles are important for refining prompts to match user expectations, but are typically excluded from the initial prompts (Kulkarni et al., 2023). Our analysis of prompts supports this observation; they usually did not include specific styles as these are added during the creation process and are not part of the initial prompt.

Learning: The AI can be adapted to specific requirements of the tasks, e.g., a specific format of the expected result that shall be produced (Reynolds & McDonell, 2021). To achieve this, examples of outcomes can be passed to the AI; currently, scientific literature distinguishes between so-called zero-shot, one-shot, and few-shot learning paradigms (Dang et al., 2022). The characteristics can be distinguished by the number of examples that are provided to the AI to fine-tune it to the expected outcome. Extant literature emphasizes the importance of providing examples to the AI to increase the outcome quality (Dang et al., 2022), however, there is also evidence that the order of examples provided changes the outcome (Wang et al., 2021).

Information Space: The information source that the AI draws on can be explicitly defined in the prompt (Liu et al., 2021). In many tasks, the need arises to search within a constrained and explicit information space for specific data. This information can be supplied either through external sources (e.g., 'search in this text for medical diagnoses') or incorporated into the prompt itself (e.g., 'answer the question drawing on the domain of sociology'), thus limiting the search to a specific domain within the model's internal information space. If the information space is not further defined, the AI will draw from its internal knowledge base. Hence, we derived the characteristics of not defined, explicit internal and explicit external information sources.

4.3. Outcome

This meta-dimension relates to the fundamental parameters that set the overall scope and expectations of the desired outcome. Drawing on the insights of Rzepka and Berger (2018), the outcome signifies the result of an interactive process. By defining the AI's learning mechanisms, and objectives, a prompt can be tailored to meet the specific needs of the intended outcome.

Chain of Thoughts: For tasks that involve multiple steps and generally have higher complexity, prompts can include the option to let the model find answers step-by-step (Shao et al., 2023). In the taxonomy, we distinguish between single-step and step-by-step prompts. Designing prompts to include step-by-step reasoning can offer several advantages. First, it enables the model to decompose high-complexity tasks into several processing steps. By computing the steps separately, additional computing power is available for steps involving high complexity. Additionally, by decomposing tasks, the user is also able to comprehend the model's problem-solving process and can intervene to improve the result if necessary (Kojima et al., 2022; Wei et al., 2022). Moreover, it should be noted that not all tasks can involve step-by-step reasoning, for instance, in the context of image generation.

Goal: Depending on the given context, prompts can have different goals. Following cognitive task analysis (Pea, 1987) and the adapted information-seeking process for cognitive strategy prompts (Schleith et al., 2022), the goals of prompts can be classified as having one of the following five characteristics: learn, look up, investigate, monitor and extract, and decide and create. Prompts to learn involve the exploration of an unknown domain. For instance, these prompts can be requests to explain a new concept in simple terms. Prompts that aim to look up information are more specific and aim to identify a specific piece of information or refine existing information. For instance, Zamfirescu-Pereira et al. (2023), highlight that senior software developers are more proficient in formulating their needs than rather new software developers. Following this example, while newer software developers need to learn more about a specific topic first, more senior software developers can likely formulate their information-seeking request more specifically. Prompts to investigate focus on analyzing or comparing known information. Monitor and extract prompts involve extracting information from a specific information source as well as monitoring information in a specific domain (Pea, 1987). Finally, prompts can be designed to create actions, i.e., make decisions or create new information of any kind (e.g., images, Oppenlaender 2022 or code, Wang et al., 2022).

5. Application of the Taxonomy

The primary objective of this taxonomy is to categorize AI prompts according to their distinct design dimensions. In the following, we will showcase the utility of the taxonomy by classifying the three examples provided in the background section (see Table 1). The first prompt (#1) is a direct request to explain an unknown concept. The input and output types are text. The interaction isn't specified; hence, it can be categorized as a computer-in-the-loop (i.e., the computer awaits human input). The prompt doesn't incorporate any role. Moreover, no learning example is included (zero-shot). The prompt does include a style (simple). The information space is not defined because no specific knowledge base is mentioned. The prompt doesn't request any chain of thoughts; hence, it can be classified as a single step. The goal is to learn more about a specified concept. The second prompt (#2) also has text as the input and output types. The interaction type is classified as human-in-the-loop because the AI takes the initiative and asks human questions. The role is defined (English teacher), and a learning example is included (one-shot). No style is defined. The information space is explicit internal because the AI processes the English phrases of the human. Moreover, the AI will conduct step-by-step reasoning (chain-of-thought). The goal can be categorized as a lookup because the information is refined. The third prompt (#3) has the input type as text and the output type as an image. The interaction type isn't specified and, therefore, it is computer-in-the-loop. No role is defined, and no learning example is specified. The style is defined through several different keywords (e.g., 'full of color, unreal engine 5'). The information space is not defined. No reasoning process is involved. The goal can be classified as create.

These three examples were selected because they depict the different types of AI prompts frequently observed during the classification process. Our taxonomy can classify and distinguish among all three prompts, indicating that it applies to a wide array of prompts. While the first and second prompts aim to extract and refine texts, the third prompt is used to generate an image. We show that the built taxonomy can distinguish the prompts by their interaction, context, and outcome.

6. Discussion

Our taxonomy describes three design components that should be addressed when creating prompts: interaction (the design of communication between humans and AI), context (the adaptation of the AI to the desired environment and requirements), and the outcome (the definition of boundaries and expected shape of results). By dissecting and organizing these components, it serves as a beneficial resource for an initial understanding of the possibilities in AI prompt design and its implications for human-AI interaction.

6.1. Theoretical Contributions

We make two contributions to IS research. First, our taxonomy offers a consolidated terminology for describing the elements of AI prompts, thereby making implicit knowledge accessible for future research. Examining prompts at varying levels allows for the exploration of the diverse dimensions of prompt shaping and their impact on AI behavior and results. The taxonomy tries to provide initial guidance on the discrepancy between AI prompts in research and practical applications. For example, we observe a strong focus of research on the learning component in the meta-dimension of context. In contrast, most analyzed prompts included a role to change the AI's behavior and adapt it to a specific context. Recognizing this disparity encourages researchers to explore and bridge the gap between theoretical discussions and practical prompt design, facilitating the translation of research findings into real-world AI systems.

Second, by investigating prompts from the perspective of IS research, and more specifically, human-AI interaction (Rzepka & Berger, 2018), we contribute to opening up the discourse on the importance of this phenomenon for the IS community and research. Our taxonomy can serve as a starting point for future research to investigate the impact of prompt design and how it changes our interaction with AI-based systems. Recent scientific literature on AI prompts is still scarce and mostly located in the research domain of computer science; hence, it mostly analyzes aspects of prompt design with a technical focus such as in-context learning (e.g., zero-shot vs. few-shot) (Dang et al., 2022; Liu et al., 2021).

Additionally, our taxonomy is not restricted to domains, offering the potential for extensions and adaptations to various fields, including text-to-image generation or software development-related prompts. This flexibility enables researchers and practitioners to apply and adapt the taxonomy across a broad range of domains.

6.2. Practical Implications

The taxonomy carries practical implications for the design and implementation of AI prompts by end-users, prompt designers, and businesses. By providing a structured blueprint for understanding the elements of prompts, the taxonomy aids in avoiding a trial-and-error process and informs design decisions during the crafting of prompts. Through consideration of the context component, prompt designers can tailor AI systems to various environments, domains, and user requirements. For example, this can be beneficial for adapting AI to business contexts, as it allows prompts to be customized to companies' specific needs. The taxonomy offers a structured overview of the possibilities to design a prompt to fit a specific department or job role through various design elements such as roles, learning examples, and styles. Moreover, the taxonomy supports end-users in creating and formulating AI prompts. As indicated by Zamfirescu-Pereira et al. (2023), lay users often fail to formulate their AI prompts effectively due to their inability to grasp the underlying problem. Finally, our taxonomy can also serve as a starting point for filtering and classifying prompts in prompt knowledge bases and help users identify suited prompts for their use cases.

Table 3. Future research agenda.

Meta-Dimension	Synthesized Literature	Exemplary Research Questions
Interaction	Research shows that AI prompt design is notably influenced by output type (Ramesh et al., 2021). Models generating images differ in prompts from text generators. Combining input types, like images with text, improves outcomes, provided congruence (Qiao et al., 2022).	• How does the output type impact the design of AI prompts? • In what contexts do combined input types prove to be beneficial?
Context	Tailoring the AI to the context via styles and roles is vital for enhancing AI-generated outcomes. However, roles remain insufficiently explored. Additionally, furnishing AI with learning examples enhances quality; however, the sequence of these examples might impact this quality (Wang et al., 2021).	• How do roles impact AI results? • When should styles be used in comparison to roles to adapt an AI to a specific task? • How does the quantity and order of provided examples impact AI results?
Outcome	The ability to formulate the goal specifically is critical to achieve the desired AI-generated outcome (Zamfirescu-Pereira et al., 2023). Moreover, when solving complex problems, the AI can decompose the problem to achieve better outcomes and make the problem solving process more comprehensible for humans (Shao et al., 2023).	• How can humans be supported at translating their goal to a suited prompt? • How are different types of reasoning perceived by humans? • Can reasoning support human error detection in factually wrong AI outcomes?

6.3. Future Research Agenda and Limitations

Drawing upon our conducted literature review and the resulting taxonomy, we synthesized the current state of literature and derived research questions (see Table 3) that are of interest to the IS community. The literature is grouped into the three meta-dimensions of interaction, context, and outcome to cluster similar research questions.

Our work is subject to several limitations. Despite our efforts to follow the proposed taxonomy development steps of Nickerson et al. (2013) and Kundisch et al. (2021) carefully and including validated conceptual foundations (e.g., the concepts of human-AI interaction), developing taxonomies and creating dimensions inductively is a subjective process. Future research could validate and potentially refine the taxonomy based on future publications and insights. Another limitation is the sheer unlimited number of prompts that could theoretically be analyzed. In contrast to other taxonomies, where the number of objects to be classified is limited, e.g., companies or business models, prompts are created within a very short amount of time. Even though we tried to reduce the arbitrariness of analyzed prompts by crawling prompt databases, we cannot rule out that we missed important prompts. Similarly, future research could systematize the analysis of prompts once there are more established and curated databases for AI prompts. Last, our taxonomy is purposely kept generic to make it applicable to a wide range of AI applications. While this can serve as a starting point for classifying prompts, it might not be detailed enough for specific purposes. For instance, following the works of Oppenlaender (2022), the style of AI prompts can be further differentiated from text-to-image prompts. Hence, future research could build upon our findings and create more specialized taxonomies for AI prompts in a specific domain.

7. Conclusion

In conclusion, this paper contributes a comprehensive taxonomy for AI prompts, offering both theoretical insight and practical guidance. The taxonomy, which dissects prompts into components of interaction, context, and expected outcome, improves our understanding of AI prompts. It facilitates a more systematic approach to prompt design, assisting users, designers, and companies in tailoring AI systems to specific needs. The practical applications of the taxonomy extend to the development of prompt knowledge bases and the selection of prompts for unique use cases. Future research should focus on expanding this taxonomy, applying it across diverse domains, and further investigating the bridge between theoretical constructs and practical applications to harness the transformative potential of AI technology.

8. References

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., … Liang, P. (2021). *On the Opportunities and Risks of Foundation Models*.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). *Language Models are Few-Shot Learners*. http://arxiv.org/abs/2005.14165

Dang, H., Mecke, L., Lehmann, F., Goller, S., & Buschek, D. (2022). *How to Prompt? Opportunities and Challenges of Zero- and Few-Shot Learning for Human-AI Interaction in Creative Applications of Generative Models*. http://arxiv.org/abs/2209.01390

Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S., & Ebel, P. (2019). *The Future of Human-AI Collaboration: A Taxonomy of Design Knowledge for Hybrid Intelligence Systems*. https://doi.org/10.24251/HICSS.2019.034

Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid Intelligence. *Business & Information Systems Engineering*, *61*(5), 637–643.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., … Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642.

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*.

Guo, Y., Liu, Y., Zhou, T., Xu, L., & Zhang, Q. (2023). An automatic music generation and evaluation method based on transfer learning. *PLOS ONE*, *18*(5).

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, *349*(6245), 261–266. https://doi.org/10.1126/science.aaa8685

Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2022). *Large Language Models are Zero-Shot Reasoners*. http://arxiv.org/abs/2205.11916

Kulkarni, C., Druga, S., Chang, M., Fiannaca, A., Cai, C., & Terry, M. (2023). *A Word is Worth a Thousand Pictures: Prompts as AI Design Material*.

Kundisch, D., Muntermann, J., Oberländer, A. M., Rau, D., Röglinger, M., Schoormann, T., & Szopinski, D. (2022). An Update for Taxonomy Designers.

Business & Information Systems Engineering, *64*(4), 421–439.
Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2021). *Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing*. http://arxiv.org/abs/2107.13586
Lubars, B., & Tan, C. (2019). Ask not what AI can do, but what AI should do: Towards a framework of task delegability. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d\textquotesingle Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 32). Curran Associates, Inc.
Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., & Söllner, M. (2019). AI-Based Digital Assistants. *Business & Information Systems Engineering*, *61*(4), 535–544.
Mayring, P. (2015). *Qualitative Content Analysis: Theoretical Background and Procedures* (pp. 365–380). https://doi.org/10.1007/978-94-017-9181-6_13
Myers, B., Hudson, S. E., & Pausch, R. (2000). Past, present, and future of user interface software tools. *ACM Transactions on Computer-Human Interaction*, *7*(1), 3–28. https://doi.org/10.1145/344949.344959
Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, *22*(3), 336–359.
Oppenlaender, J. (2022). *A Taxonomy of Prompt Modifiers for Text-To-Image Generation*. http://arxiv.org/abs/2204.13988
Pea, R. (1987). User Centred System Design-New Perspectives on Human/Computer Interaction. *J Educ Comput Res*, *3*.
PromptHero. (2023). *ChatGPT prompt: I want you to act as a social media influencer - PromptHero*. https://prompthero.com/prompt/f0e3c1c8d70-chatgpt-social-media-influencer
Qiao, H., Liu, V., & Chilton, L. (2022). Initial Images: Using Image Prompts to Improve Subject Representation in Multimodal AI Generated Art. *Creativity and Cognition*, 15–28.
Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M., & Sutskever, I. (2021). Zero-Shot Text-to-Image Generation. In M. Meila & T. Zhang (Eds.), *Proceedings of the 38th International Conference on Machine Learning* (Vol. 139, pp. 8821–8831). PMLR.
Reynolds, L., & McDonell, K. (2021). *Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm*. http://arxiv.org/abs/2102.07350
Rzepka, C., & Berger, B. (2018, December 14). User Interaction with AI-enabled Systems: A systematic review of IS research. *ICIS 2018 Proceedings. 7.*
Schleith, J., Norkute, M., Mikhail, M., & Tsar, D. (2022). Cognitive Strategy Prompts. *Creativity and Cognition*, 29–37.
Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G.-J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information & Management*, *57*(2), 103174.
Seeber, I., Waizenegger, L., Seidel, S., Morana, S., Benbasat, I., & Lowry, P. B. (2020). Collaborating with technology-based autonomous agents. *Internet Research*, *30*(1), 1–18.
Shao, Z., Gong, Y., Shen, Y., Huang, M., Duan, N., & Chen, W. (2023). *Synthetic Prompting: Generating Chain-of-Thought Demonstrations for Large Language Models*. http://arxiv.org/abs/2302.00618
Sorensen, T., Robinson, J., Rytting, C. M., Shaw, A. G., Rogers, K. J., Delorey, A. P., Khalil, M., Fulda, N., & Wingate, D. (2022). *An Information-theoretic Approach to Prompt Engineering Without Ground Truth Labels*.
Terveen, L. G. (1995). Overview of human-computer collaboration. *Knowledge-Based Systems*, *8*(2–3), 67–81.
Wang, A. Y., Wang, D., Drozdal, J., Muller, M., Park, S., Weisz, J. D., Liu, X., Wu, L., & Dugan, C. (2022). Documentation Matters: Human-Centered AI System to Assist Data Science Code Documentation in Computational Notebooks. *ACM Transactions on Computer-Human Interaction*, *29*(2), 1–33.
Wang, Y., Yao, Q., Kwok, J. T., & Ni, L. M. (2021). Generalizing from a Few Examples. *ACM Computing Surveys*, *53*(3), 1–34.
Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26*(2), xiii–xxiii.
Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2022). *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models*. https://doi.org/2201.11903
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*. http://arxiv.org/abs/2302.11382
Wiethof, C., & Bittner, E. A. C. (2021). Hybrid Intelligence – Combining the Human in the Loop with the Computer in the Loop : A Systematic Literature Review. *Proceedings of the 42nd International Conference on Information Systems (ICIS 2021), October*, 0–17.
Wu, T., Terry, M., & Cai, C. J. (2022). AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts. *CHI Conference on Human Factors in Computing Systems*, 1–22.
Yan, W., Zhang, Y., Abbeel, P., & Srinivas, A. (2021). *VideoGPT: Video Generation using VQ-VAE and Transformers*. https://doi.org/2104.10157
Zamfirescu-Pereira, J. D., Wong, R. Y., Hartmann, B., & Yang, Q. (2023). Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–21.