Empower students' autonomy in design process when using Chat GPT for their academic tasks

A Master Graduation Thesis by
Hui-Ru, Yang

TU-Delft, November 2024

Executive Summary

Empowering Student Autonomy in AI-Supported Learning

The integration of artificial intelligence (AI) tools, such as ChatGPT, into higher education has reshaped how students approach academic tasks. While these tools offer enhanced efficiency and convenience, they have also raised significant concerns about their potential to undermine students' autonomy, critical thinking, and decision-making abilities. This design thesis seeks to address these concerns by exploring how AI systems can better empower students, focusing on autonomy values such as control, decision-making, and engagement.

To address the research problem, a structured, multi-phase approach based on VSD was employed. This approach focused on identifying user needs, defining design goals, and iteratively testing conceptual solutions to develop a user-centered design. This study identified key challenges and opportunities in how design students interact with ChatGPT for tasks like brainstorming, writing, and research through literature reviews, interviews, workshops, and evaluations. The process emphasized empowering students to take an active role in their learning processes, mitigating overreliance on AI, and fostering meaningful engagement.

Key Findings

1. Aligning with Users' Real Intentions: Features such as Off-Topic adjustment and CardNote excel in helping users focus on key topics, streamline information, and achieve clarity. However, high-complexity scenarios highlight gaps in scalability and efficiency.
2. Avoiding Misleading Communication: Dictionary and CardNote significantly reduce ambiguity by clarifying terminology and ensuring conversation continuity. Conversely, the Prompt Guide feature lacks specificity, leading to frustration in advanced use cases.
3. User-Centered Adjustability: Interactive customization tools, like Off-Topic and CardNote, empower users to adjust AI outputs effectively. However, reliance on manual input creates inefficiencies when managing large volumes of information.

Design Proposals

Four conceptual features were designed and evaluated:
1. Off-Topic: Enables users to refocus discussions by filtering irrelevant content, fostering smoother interactions.
2. Prompt Guide: Supports idea generation by offering structured prompts, though improvements are needed for precise user guidance.
3. Dictionary: Clarifies technical terms, improving communication and reducing cognitive load.
4. CardNote: Enhances information management by allowing users to save, categorize, and revisit conversation highlights.

Research Implications

This study underscores the importance of aligning AI tools with autonomy-related values in education. By empowering users to make decisions, refine AI outputs, and maintain control over their interactions, AI can transition from passive assistants to active Learning collaborators. However, future designs must address limitations such as contextual adaptability, scalability, and support for cumulative discussions to achieve this vision.

Next Steps

Future iterations will focus on:
- Incorporating automated content prioritization to manage high-information scenarios.
- Enhancing context-aware functionalities to deliver tailored suggestions and explanations.
- Improving visual feedback mechanisms to reinforce usability and engagement.

This thesis lays the groundwork for creating AI systems that assist and inspire students to take ownership of their academic journeys, fostering a culture of proactive Learning and creative exploration.

Author

Hui-Ru, Yang
MSc. Design for Interaction
Faculty of Industrial Design Engineering (IDE)
Delft University of Technology

Supervisory team

Chair
Dr. Sara Colombo
Assistant Professor
Department of Sustainable Design (SDE) department

Mentor
MSc. Eric Gu
PhD candidate
Design, Organization, and Strategy (DOS) department

Exploring Design Opportunity CH1 CH2

Objective

- Research current application and research gap on AIED.
- Research current theory (VSD) and research gap on Ethic/ Autonomy.
- Exploring opportunities for design intervention.

I: Contextualization CH3 CH4

Objective

- Research current application on AIED and definition on Autonomy.
- Understand current context from students' viewpoint.
- Identify the needs and pain transferred from Autonomy perspective.
- Identify opportunities for design intervention.

II: Design Goal CH5

Objective

- Define Design Goal
- Define Design criteria

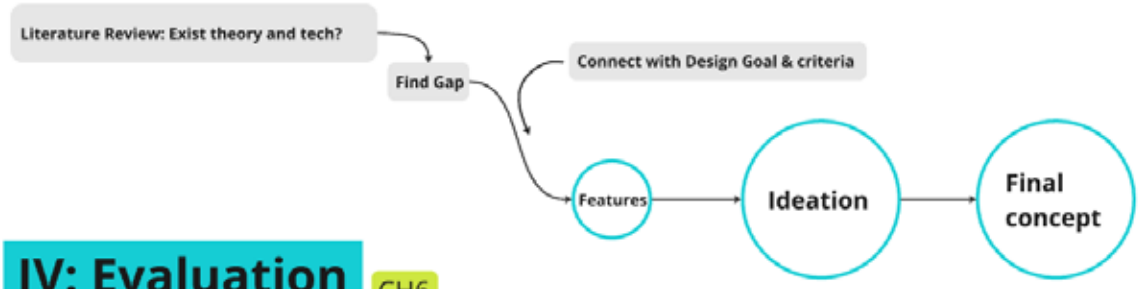
Main Goal

Promote a user experience that enables humans to communicate their ideas more effectively to an ideation AI system, ensuring a productive and intuitive human-AI collaboration.

III: Conceptualization CH5

Objective

- Research for current study and technique again.
- Define the problem statement
- Develop the features based on Design Goal & Design criteria & theory



IV: Evaluation CH6

Objective

- To evaluate whether the final concept proposal effectively addresses the design goal.

V: Recommendation & Conclusion CH7

Objective

- To recommend further iterations and possible developments for the concept

Table of Content

Executive Summary
Report Structure

CH1 Introduction 7
1.1 Research Background 8
1.2. Research Problem 8
1.3 Research Theme 9
1.4 Research Objective 9
1.5 Research Significance 9
1.6 Out of Scope 9

CH2 Literature Review 10
2.1 Education for AI" and "AI for Education." 11
2.2 The Value Sensitive Design Approach 12

CH3 Context analysis14 14
3.1 Societal Challenges 15
3.2 Defining Autonomy 16

CH4 Contextualization & Value analysis 18
4.1 Interview session: Stakeholder analysis 19
4.2 Interview Result 19
4.3 Interview Findings 21
4.4 Workshop session: Laddering Stakeholder's Deeper Value and Needs 22
4.5 Workshop Result: Value and norm identification 24
4.6 Values and needs' Analysis process 25

CH5 Design 31
5.1 Problem statement 32
5.2 Design Directions 32
5.3 Process of conceptualization: 2 Round 34
5.5 Final Design Concept Proposal 39

CH6 Evaluation 46
6.1 Evaluation plan 47
6.2 Results and Discussion 48
6.3 Summary of Findings 59

CH7 Recommendation & Conclusion 61
7.1 Future Recommendation 62
7.2 Limitation 63

References 64
Appendix 67
Appendix A: Project Brief 68
Appendix B: Workshop analysis- All broken down needs and pains 74
Appendix C: Code book 76

1 Introduction

CH1. Introduction

1.1. Research Background

Artificial Intelligence (AI) agents are increasingly used in various educational scenarios to interact with humans. Chatbots and related algorithms can simulate human interactions and generate human-like text based on natural language input (Adiguzel, 2023). For instance, students use ChatGPT for a range of academic tasks, including coding guidance, writing support, and idea conceptualization. Additionally, ChatGPT serves as a virtual tutor for instructors, offering one-on-one communication capabilities, in tasks such as generating course materials and providing suggestions that effectively address individual students' specific queries.

However, the convenience offered by such technology has led to user's overreliance, raising concerns about the potential erosion of human autonomy. Floridi and Cowls (2019) discussed the concept of autonomy, highlighting concerns that "the growth in artificial autonomy may undermine the flourishing of human autonomy" (Floridi & Cowls, 2019, p. 7). While humans believe that using AI in decision-making can be beneficial and timesaving, it can also diminish the brain's natural cognitive processes, thereby reducing cognitive capabilities (Jarrahi, 2018). Research indicates that using AI models for academic writing might reduce critical thinking abilities and innovative thought processes among students and teachers (Ahmad et al., 2023). This study aims to explore how to enhance students' autonomy when using ChatGPT for academic tasks through the application of the Value-Sensitive Design (VSD) framework.

1.2. Research Problem

This study aims to research how AI models impact the autonomy of college students aged 25 to 30 in their learning experiences, particularly in terms of interaction and shifts in mindset and behavior. Additionally, it will explore how the Value-Sensitive Design approach can empower AI models like ChatGPT to enhance students' autonomy and foster a supportive learning environment.

The specific research objectives are as follows:

1. Determine the changes in mindset and behavior of stakeholders when interacting with ChatGPT for academic tasks.
2. Determine the autonomy values of stakeholders in education when using ChatGPT for their academic tasks.
3. Explore how the Value-Sensitive Design approach can enhance the autonomy values of AI models.

These objectives aim to provide insights into how AI models can promote student autonomy in an educational context. Ultimately, the study seeks to develop a prototype for new features to be embed in future designs of AI agents in higher education.

1.3 Research Theme

- How might we empower the students' **autonomy** values to encourage **proactive engagement** in process when using Chat GPT for their academic tasks?

1.4 Research Objective

- To enhance student autonomy, ensuring they have greater **control and decision-making power** when using ChatGPT for academic tasks.

1.5 Research Significance

Enhancing student autonomy is crucial for improving learning outcomes and long-term development. Autonomous learners are more likely to maintain high levels of engagement and motivation, leading to deeper understanding and application of knowledge. Furthermore, fostering student autonomy equips them with the adaptability and continuous learning skills necessary in a rapidly changing society. This is vital for their personal growth and career development. Therefore, this study not only aims to improve educational quality but also emphasizes ethical considerations to ensure the responsible use of AI technology, ultimately contributing to sustainable educational advancements.

1.6 Out of Scope

To ensure the clarity of the research objectives, the study will focus on the issue of "students' lack of autonomy in the academical process" rather than on "critical thinking skills level" or "learning path." Specifically, the research will concentrate on enhancing students' autonomy by ensuring they have more control and decision-making power when using ChatGPT for academic tasks. The goal is to guide students to actively participate in the learning process, rather than passively relying on ChatGPT-generated answers.

During this process, students' engagement and reflection in learning might also foster the development of critical or diverse thinking skills. While the development of these skills is important and related, it is not the primary objective of this research. **These skills are considered incidental benefits that support the enhancement of student autonomy.**

In summary, this thesis will focus on how to design a system to enhance student autonomy related to control and decision-making power instead of discussing the quality of ideas or the degree and tendency of critical thinking.

2

Literature Review

CH2 Literature Review

2.1. Education for AI" and "AI for Education."

To ensure clarity and precision in subsequent discussions, it is essential to distinguish key terms between "Education for AI" and "AI for Education" and their categories. We refer to Schiff's (2022) classification for five primary themes: three themes relate to the concept of "Education for AI," understood as training, and two themes are associated with "AI for Education" (AIED).

Generally, "Education for AI" focuses on equipping individuals with AI skills and knowledge to meet future labor market demands and technological challenges. This incorporates (1) training AI experts, (2) preparing the workforce for AI, and (3) promoting public AI literacy. Training AI experts involves developing future practitioners such as computer scientists and engineers, as seen in initiatives like Singapore's AI Apprenticeship Program (Smart Nation and Digital Government Office, Singapore 2019). and Lithuania's AI Master's programs (Minis- try of the Economy and Innovation 2019). Preparing the workforce for AI includes educational and training programs aimed at helping workers adapt to AI-driven changes in the labor market, emphasizing lifelong learning and curriculum reform. Public AI literacy aims to increase general knowledge and engagement with AI.

In contrast, "AI for Education" refers to the application of AI technologies to enhance educational processes, including (1) teaching, learning, and (2) administrative tasks. Teaching and learning involve AI-based educational tools such as intelligent tutoring systems, pedagogical agents, and predictive assessments. Countries like China, India, and Spain have incorporated these tools to personalize education and support students with special needs. Administrative tools involve using AI to manage educational systems, including admissions, promotions, and graduation decisions. For instance, India has proposed automating teacher hiring and transfer systems (NITI Aayog, India 2018) , while Malta is exploring AI for scheduling and facility management. (Office of the Prime Minister, Malta 2019).

Despite the significant global focus on AI ethics, there is relatively little attention to the ethics of AI in education (AIED). This divergence arises primarily due to three main reasons. Firstly, policy priorities tend to emphasize economic growth and national competitiveness, often overlooking the specific ethical issues related to AIED. Additionally, most policies focus on developing AI talent and preparing the workforce rather than applying AI to improve education. The immediate need for AI-skilled professionals to meet market demands leads to a priority on rapid training programs. From a long-term strategic perspective, policymakers view developing AI talent and workforce transformation as crucial for sustained economic benefits. And concerning the limited government resources, there is a tendency to allocate funds to areas with direct economic returns, rather than investing in AI applications in education, which may require longer-term observation and investment.

Secondly, policymakers and the public lack sufficient understanding of AI applications in education and their potential impacts, leading to inadequate ethical discussions. Thirdly, the implementation and research on AI in education are relatively new, resulting in a lack of comprehensive ethical considerations.

Above all, researching the ethics of AI and its applications in education is essential to ensure the responsible use of these technologies. While AI becomes increasingly integrated into educational systems, understanding its ethical implications and practical challenges will help formulate more balanced and comprehensive applications, ensuring that technological applications not only drive economic development but also genuinely enhance the ethical side of education. Thus, the study will scope in the "teaching, learning" sector in AIED.

2.2 The Value Sensitive Design Approach

Value Sensitive Design (VSD) is a suitable method for exploring ethics in technology design because it emphasizes integrating human values directly into the design process. It defines value as "what a person or group of people consider important in life" (Friedman et al., 2013, p. 2). The identification of values in VSD depends on the specific systems, contexts, stakeholders, and application domains under analysis. VSD comprises three iterative and integrative investigations: conceptual, empirical, and technical. VSD encourages designers to continually reflect on and revise the ethical dimensions of their work. This iterative approach allows for ongoing assessment and incorporation of feedback, ensuring that the technology remains aligned with ethical standards throughout its lifecycle.

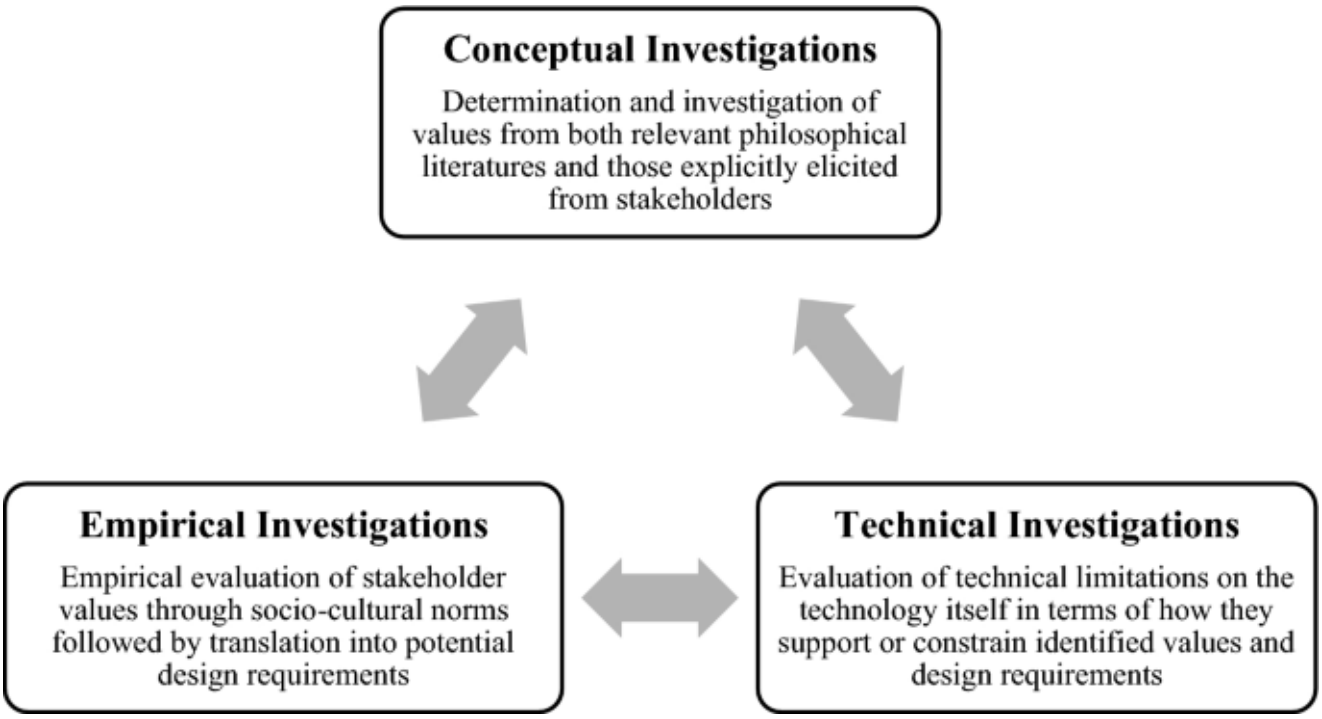


Fig 2.1.1 Value Sensitive Design (VSD) method

2.2.1 Values Hierarchy

The 'values hierarchy' developed by van de Poel (2013) providing a systematic approach to translating abstract values into concrete design requirements. This hierarchy consists of three layers: (1) values, which are general principles to be promoted for their own sake; (2) norms, which are guidelines or prescriptions for action; and (3) design requirements, which are specific criteria to be met as fully as possible. These hierarchies can be constructed both top-down and bottom-up. The structure ensures that practitioners address the fundamental points proposed in the VSD framework. Furthermore, this tool helps ensure that design decisions are aligned with both ethical and practical considerations, facilitating a deeper understanding and realization of values in the technology development process. By structuring the translation of values into actionable design elements, the values hierarchy enhances the social acceptability and ethical compliance of the resulting technology. The study will specifically explore the norms related to the value of autonomy.

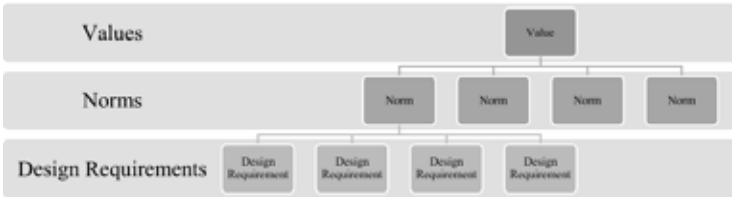


Fig 2.2.1.1 Values Hierarchy

2.2.2 AI4SG-VSD Process

The VSD approach has been increasingly adopted and extended into various domains. In the field of artificial intelligence (AI), Steven Umbrello and Ibo van de Poel (2021) have proposed a framework based on VSD to address the unique challenges posed by AI technologies. This framework consists of four main processes: context analysis, value identification, formulating design requirements, and prototyping. For the sake of brevity, this paper does not explain the definitions of these four stages.
Umbrello and van de Poel (2021) argue that AI4SG factors function like norms, aiming to ‘maximize’ or ‘minimize’ certain values or design requirements. This helps bridge the gap between abstract values and concrete design requirements.
The objective of this paper is to apply the AI4SG methodology, using VSD, to propose a design for ChatGPT that enhances student autonomy.

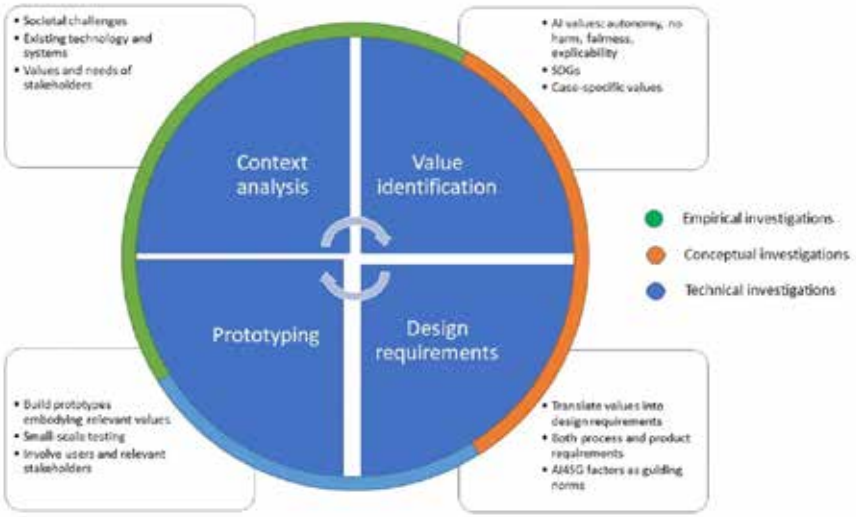


Fig 2.2.2.1 VSD design process for AI technologies

3

Context analysis

CH3 Context analysis

3.1 Societal Challenges

The integration of Artificial Intelligence (AI) in education presents several societal challenges, particularly concerning student autonomy. One significant issue is the potential for AI to foster a sense of dependency among students. Studies have shown that students may increasingly rely on AI to perform academic tasks, leading to a decline in their motivation to engage deeply with the learning material and to develop critical thinking skills. This reliance can result in a form of intellectual laziness, where students prefer to utilize AI-generated information rather than exerting effort to understand and analyze content independently.

Moreover, the convenience and efficiency provided by AI technologies can inadvertently undermine students' decision-making abilities. As AI systems become more sophisticated and capable of handling complex tasks with minimal human intervention, there is a growing risk that students will become passive recipients of knowledge rather than active participants in their learning processes. This shift could lead to a reduction in their capacity for autonomous decision-making and critical evaluation, which are essential skills for personal and professional development in a rapidly changing world.

The societal challenge, therefore, lies in balancing the benefits of AI in enhancing educational efficiency and accessibility with the need to preserve and promote student autonomy. This involves designing AI systems that support students in making independent decisions and developing their critical thinking abilities rather than replacing these cognitive functions. By addressing this challenge, educators and policymakers can ensure that AI contributes positively to educational outcomes without compromising the development of essential life skills.

In summary, while AI has the potential to revolutionize education by improving teaching quality and efficiency, it is crucial to address the ethical implications related to student autonomy. Ensuring that AI applications in education empower students to take control of their learning and make informed decisions is key to leveraging the full potential of AI while safeguarding the integrity of the educational process.

3.2 Defining Autonomy

The concept of human autonomy is an essential element in ethics. In General, autonomy is defined as willingness and self-endorsement of one's behaviors, which external forces should not restrict. In the following paragraph, we will lay out a few conceptual definitions to clarify the notion as we employ it in this thesis.

From the perspective of Philosophical Positions, John Stuart Mill mentioned the concept of "self-determination" in his famous work On Liberty in the early time, which is close to the term "autonomy" nowadays, broadly as the capacity to be as itself, live with its reasons and motives that are determined as one's own and without the product of manipulative or distorting external forces (Christman, 2018). Additionally, Mill mentioned that the value of this capacity is not limited to any domain but is a characteristic that can apply to any aspect of an individual's life. (Christman 1989, 2018).

Although Individuals should have the right and capacity for rational autonomy, Hill (2013) recognized that external forces like physical force, coercive threats, deception, manipulation, and oppressive ideologies can limit our ability to think and act freely. The notion of "hindrance to autonomy" is a crucial concept relevant within specific technological contexts, and we will illustrate this in detail later in the three directions.

When mentioning the notion of "hindrance to autonomy, it is essential to note heteronomy, the opposite of autonomy. Heteronomy means that one's actions are influenced by internal or external pressures perceived as controlling (Ryan & Deci, 2017). Feeling controlled can be either direct or minor and imperceptible, like when technology forces someone to perform undesired actions that she does not value (e.g., clicking through unwanted online pages infringing on one's minor autonomy).

Besides the external factors, the internal pressures or heteronomous forces (Ryan, 1982) can also hinder autonomy. For instance, technology users may develop compulsions leading to overuse, often observed with video games and social media (e.g., Przybylski et al. 2009). Przybylski et al. (2013) discovered that people who frequently used social media experienced higher levels of FOMO (fear of missing out), linked to reduced satisfaction with basic needs, decreased autonomy, and lower mood. These examples suggest that, even if users choose technology willingly, they may nonetheless experience control without being fully autonomous (Ryan & Deci, 2017). Therefore, this research will consider the contexts in how to help people regain control over their use of AI technologies.

To help recognize the experience easily, these points are combined and categorized into the STA paradigm, sense, think, and act (Siegel, 2003; Hayles, 2005), which includes a specific problem-solving task, into sense autonomy, thought autonomy and action autonomy to capture critical characteristics for understanding autonomy in AI and technology contexts.

Specifically,

- Sensing autonomy refers to humans' feelings (e.g., willingness, volition, and endorsement) and emotions (e.g., emotions, flow, tiredness) to sense the surrounding environment when using AI.
- Thought autonomy refers to the ability of a human to think, reflect on their actions, and make decisions or plans independently.
- Action autonomy refers to the ability of humans to act, change behaviors, execute plans, or perform certain behavioral activities independently.

While this is not a complete or sufficient conceptualized category for defining human autonomy in AI systems, it offers a valuable foundation that provides a basis for recognizing a large number of the autonomy tensions that arise within these contexts, which will be demonstrated within the workshop session and discussion in Chapter 5, 6. It is worth noting that artificial autonomy is not dichotomous but varies along a continuum from no autonomy to full autonomy (Beer, Fisk, & Rogers, 2014). Similarly, human autonomy is also not dichotomous.

4

Contextualization & Value analysis

CH4 Contextualization & Value analysis

4.1 Interview session: Stakeholder analysis

In this chapter, we interviewed 10 higher education students aged between 25 and 30, including master students from Design, Aerospace Engineering, Electrical Engineering, and the Mathematics and Computer Science departments.

The main objectives are to understand the following:

1. Awareness and Usage: Knowing how ChatGPT is used, the frequency of interaction, and the purposes behind its usage.

2. Personal Reflections: Positive and negative experiences shared regarding the use of ChatGPT.

3. Autonomy Considerations: The participants' understanding of the concept of autonomy, as well as the limitations and expectations encountered when using ChatGPT.

4.2 Interview Result

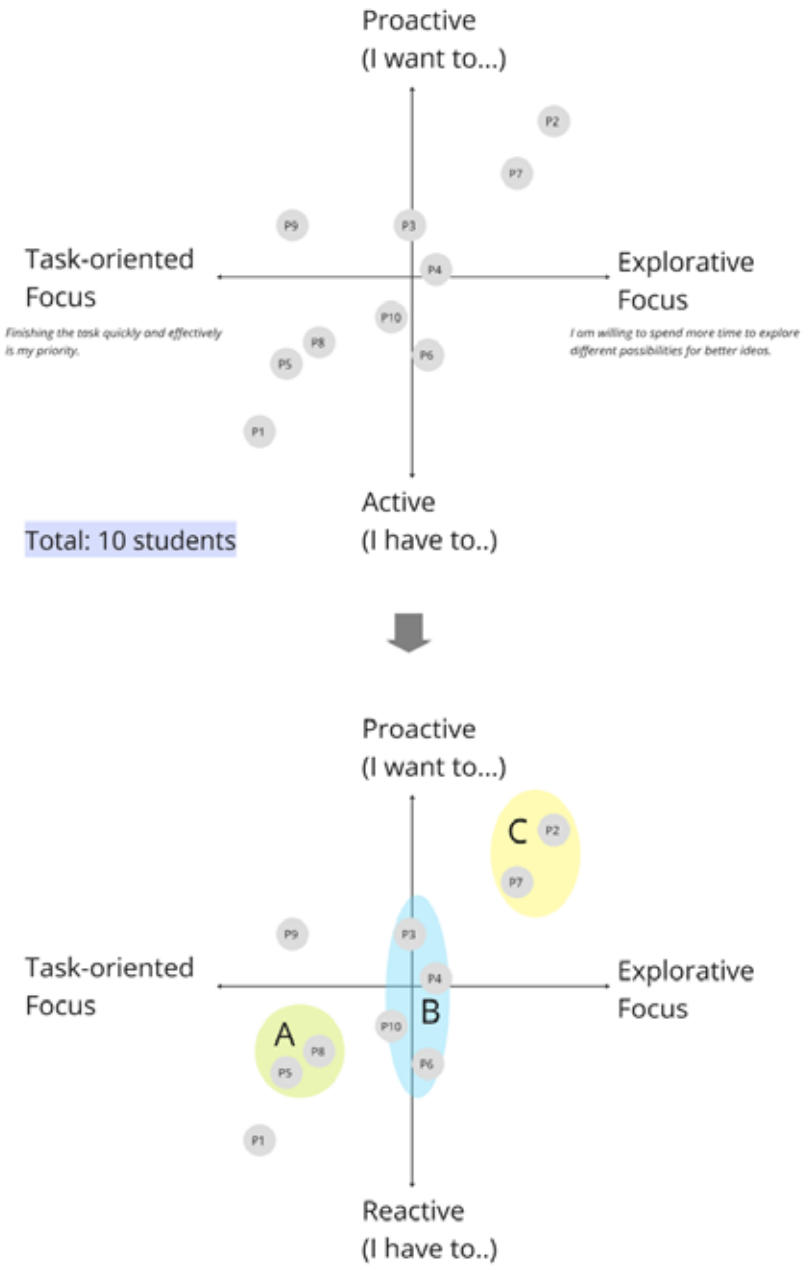


Fig 4.2.1 Interview Analysis - Using Type and Attitude

Interview Quote

Attitude types and the Quote:

TypeA: Efficiency-focused

Characteristics:

- Prioritizes speed and efficiency, with relatively low demands on the quality of results.
- Heavily relies on ChatGPT to complete tasks quickly.
- High trust in ChatGPT's output, with minimal reflection or adjustment.

P1 1. To save time, I often just copy-paste and then little adjust the text to fit what I need.
P5 2. ChatGPT is such a handy tool. When I need help with a structure or ideas, it gives me a great response right away. I typically start with that.
P8 3. Efficiency is the most important to me. For me, it's not about deadlines or time limits; I just believe every task should be done quickly and effectively so that I have more time to relax.

Type B: Hesitant and Dependent Users

Characteristics:

- High dependency on ChatGPT, but lacks confidence, often feeling hesitant during use.
- Worries that over-reliance on AI might impact their own abilities, leading to a doubtful attitude towards each use.
- Lower efficiency, prone to repeated self-doubt and revisions.

P3 1. I feel like it might change the way I think in the future, you know? It's so easy to get an answer with just a press of a button. But at the same time, I'm combining it with other tools to make the most of their strengths. Like, I use Quilbot to rephrase sentences or Researchrabbit to find relevant papers. But honestly, I do feel a bit guilty about it.
P4 2. I'm worried we might be losing some of our human abilities. Because I used to develop ideas by myself based on my knowledge, books and references instead of ChatGPT.
P6 3. Honestly, I can't always tell the difference between good and bad output. But I wish I could.
P10 4. The choice of words in ChatGPT can be really weird sometimes, you know, like using unfamiliar terms.

TypeC: High Autonomy with Balanced Focus on Quality and Efficiency

Characteristics:

- Highly autonomous, capable of adjusting strategies flexibly to achieve the best results.
- Focuses on the quality and accuracy of results, always exploring tweaking and optimizing prompts for their needs.
- Demonstrates reflectiveness when using ChatGPT, making adjustments based on actual needs.

P2 1. I usually write the assignment myself first, and then I use ChatGPT just to fix the grammar or get some advice on the structure.
P7 2. I spent a lot of time experimenting with different prompts to get better results.
P9 3. Sometimes the prompts I use don't quite capture what I'm trying to say, so I end up rephrasing them a few times.

In summary, all three types of users share the need to communicate and interact with ChatGPT. However, Type B users show the most significant design potential, as their motivation to change is the strongest among the three. In the interviews, Type B users expressed a desire to "learn how to better use ChatGPT" and "hope to enjoy the process more, making collaboration with ChatGPT a more smooth relationship."

This desire aligns closely with the direction of Type C users and could be a valuable design direction for future exploration: How to enable users to achieve a greater sense of awareness and control in their collaboration with ChatGPT.

4.3 Interview Findings

Take Away

1. Need for More Guidance: Students' understanding of the term "autonomy" is still abstract, and they lack sufficient comprehension. Despite being provided with positive and negative academic definitions before the interviews, they still lack a clear conceptualization of this term.

2. Targeting Specific Groups: During the interviews, we asked about users' experiences with ChatGPT for assignments, including academic writing, creative thinking, applying new knowledge for tasks, coding, literature search, and idea structuring. While the contexts were diverse, the professional needs varied significantly across departments. Also, we found that even within the same context, users might adopt different behaviors and mindsets. ***Therefore, to deepen the needs assessment, we decided to focus the next phase of the workshop on design students specifically.*** **The main reason is the sample size limitation, as most interviewees were from design disciplines, resulting in a higher amount of feedback related to creative thinking exploration.**

Hence, to concentrate our research, we decided to make creative thinking exploration the focus for the next phase of the workshop, targeting this specific group.

3. Abstract Emotional Feedback: The interviews followed a semi-structured format, covering approximately 30% process actions and 70% reflections. Surprisingly, many of the insights were about people's emotional reactions to using ChatGPT, suggesting we should look into this more in the next phase.

4. Design Opportunities: Among the three types, ***Type 2 is considered to have the highest design potential, as users exhibit the strongest "motivation" to make changes.***

In the next chapter, we will adopt a more emotion-, action-, and thought-oriented definition of autonomy to guide the collection of user experiences, aiming to identify and consolidate specific pain points.

4.4 Workshop session: Laddering Stakeholder's Deeper Value and Needs

Workshop: Procedure and Participants

We carried out a user study session involving 4 master students aged 25 to 30 in TU-Delft. The participants were recruited from IDE design faculties. Additionally, all of them were well-experienced and high frequency ChatGPT users for their usual academic task, fostering a more dynamic experience sharing. The study spanned a duration of 90 minutes and consisted of three sessions:

4.4.1. Session 1:

Metaphor on Autonomy (15-min)

Understanding the autonomy meaning in an educational context

In the previous interviews, we found out that the interviewees were mainly unfamiliar with the word "Autonomy". While some of them shared their thoughts insightfully, others found the concept abstract or confusing. To mitigate this, we needed alternative phrasings to understand users' sense of autonomy and agency. In the previous HCI research, the rephrasing words like "in control", or "empowered" using to illustrate the Autonomy, were proved being more understandable and capable of capturing study participants' judgments (Phil Benson 2010, Kai.et., al 2023, Janet 2007) Thus, we began by encouraging the students to share their experience with ChatGPT with the following question:

What are your feelings/ interactions when using an AI tool to support academic tasks?

We chose DIXIT Cards as the tool to lead the potential hints. The cards present an abstract drawing, the category of which is undefined, and the meaning is open for interpretation. (DIXIT Cards, 2014). The studies have shown that an unstructured source of inspiration represented herein by DIXIT Cards is more likely to help designers support exploration in the idea generation process since these cards allow for free interpretation (Kwiatkowska, Szóstek, & Lamas, 2014; Sousa, 2020). Therefore, we utilize the cards for finding the autonomy-related keywords. Each student can pick two cards that best represent their using context and share thoughts related to the cards' elements with others without judging their answers as right or wrong.

This method helped students in our study express their views on autonomy. Moreover, this also served as a great starting point for a simple warm up.

4.4.2 Session 2:

Find Value scenarios by Semi-structured Interview and laddering technique (45-min) + 10 mins break.

In this session, we carried out a semi-structured focus group interview. We began by prompting the participants the Autonomy phrasings from the last sessions and some Autonomy's definitions in Chapter 3 literature research. Next, we encouraged them to wrote down their using scenario or event with ChatGPT on the post-its based on these hints. After that, we encouraged the participants to reflect and share their experiences related to autonomy in the group using the laddering interview technique. By progressively exploring each event through this method, we aimed to uncover the specific motivations and feelings behind their actions. Laddering allowed us to delve deeper into their responses, moving from surface-level attributes to the underlying values and emotional benefits associated with their experiences. Ideally, we hoped that through sharing their experiences, they would develop a sense of connection and extend their insights to a broader range of autonomy-related experiences.

4.4.3 Session 3: Cluster ideas (20-min)

In this session, participants were introduced to three autonomy categories— "Sense, Thought, Action"—from Chapter 3. They were asked to cluster their similar and differences value scenarios with the summaries and tagged with these three categories. During tasks, participants were encouraged to collaborate as a group and sharing loud when doing the clustering and discussions. Researcher observed and noted points for discussion in the follow-up interview to clarify or expand on particular points. Following this, the complete information for background context could be well-collected.

4.5 Workshop Result: Value and norm identification

This research focuses on finding the main user needs or pain points regarding ChatGPT use in relation to autonomy values. Friedman and Hendry (2019) have proposed 14 more specific methods for finding values in Value Sensitive design approach.

Since autonomy is our research's primary object of ethical value, we use specific methods to delve deeper into this aspect. The (6) value-oriented semi-structured interview and (8) value-oriented coding manual are particularly suited for this, as they allow for a more profound exploration and identification of stakeholders' values, unlike the broadly-values methods (e.g., envisioning cards).

4.5.1 Results– session 1

Here are the words we got from four participants. Each of them picked two cards with several metaphors. (Figure XX). These words formed the basis of our quality code, highlighting the significance of input in understanding autonomy. (Table XX)

Participant	Cards	Words
P1	1	Navigating
	2	you don’ t have to be good at everything, magic, superpower
P2	3	Freedom to learn new things, Logic step by step, No ChatGPT
	4	Capability, Human Agency
P3	5	unexpected outcome from machines
	6	manipulate, uncertainty, combining different resources, witch, steering the wheel, superpower
P4	7	Black box of AI opaque
	8	Self-control, ownership, control over time, responsibility

4.5.2 Value-oriented semi-structured interview in Group sharing

The second step of our research is dedicated to identifying autonomy-related values in the context of ChatGPT use in academic tasks. In the previous step, the participants demonstrated a concrete understanding of the term 'autonomy' in this context. They were then asked to share their experiences of feeling 'forced' or experiencing 'Heteronomy', as we defined in Chapter 3.

The records were transcripted from the workshop, which covered how AI tools are used and the 'Heteronomy' experience in a broad sense. During the process, the hints and examples collected from the previous interviews are given as the prompt for better leading and discussion flow. For example, give the focus group one direct and imperceptible example from previous interviewees for each. The red post-its presented each event for the scenario coding references afterward. The focus group, which specifically explored AI-tool interaction experiences with each other, was uploaded and analyzed together through the coding approach. While one in the focus group provided the primary information regarding specific interactions with AI tools, the others were inspired by related experiences. The process provided an overall understanding of user values and background context with different participants' experiences by justifying their combined analysis and offering concrete insights into the technical and psychological aspects of AI tools and autonomy. Next, the events (red post-it) are broken down into individual heteronomy components regarding the autonomy categories—"Sense, Thought, Action"—from Chapter 3. While the broken-down process needs sensitive exploration and reflection, semi-structured group interviews are well-suited for identifying stakeholders' values.

This way triggered a value awareness when doing the group sharing. Until this step, the complete information for background context was almost well collected.

4.6 Values and needs' Analysis process

Collecting Events and Broke down Elements

The Value-oriented coding manual is a well-suited method for VSD, which approaches iterative and inductive processes after exploring and identifying stakeholders' values concerning autonomy experience used with the AI tool in the last section. The iterative and inductive can connect and cluster similar events through coding phases by three autonomy categories—"Sense, Thought, Action" and uncover the underlying values, needs, and opinions of users and other stakeholders for the whole picture.

In the first stage, the events (red Post-it) and the broken elements (green Post-it) were labeled individually without attempting to establish connections between them through open coding. (figure 1) Next, the three autonomy categories—"do, think, sense" are recognized and labeled in the coding process by clustering them. In this step, we can see the clear unit events and elements from three clusters.

Finally, the context statements from these unit elements were built based on the event discussed in semi-structured interviews done in precious Group sharing. This round of code categories by Type ("Sense, Thought, Action"), Events, Breakdown Elements, Topic (Context statement), and Frequency of Mentions can be exported in the following process.

Fig 4.6.1 User analysis (More Details in Appendix B)

Step 1:

We collected all the quotes from the four participants in the workshop (six example sections shown below). By adopting a semi-interview combined with workshop format, we were able to elicit related experiences that resonated well among participants, sparking further experience sharing and discussion. We recorded all the quotes related to the loss of autonomy and gathered similar Post-it notes together.

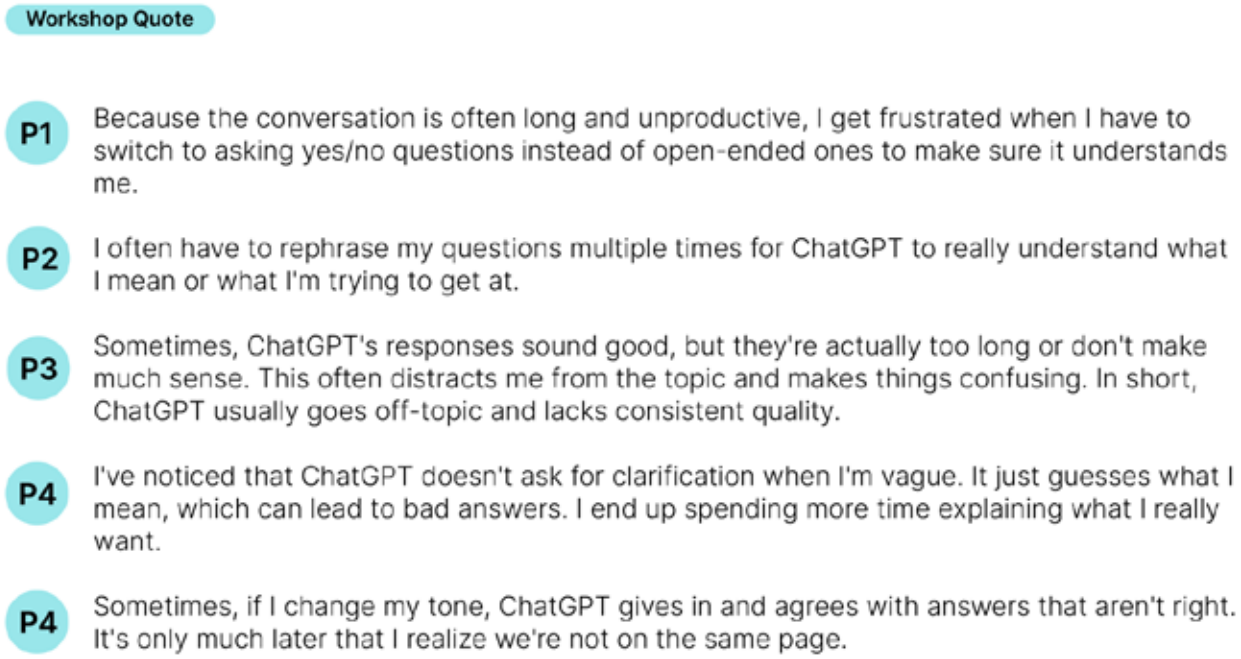


Fig 4.6.2 Quotes from the workshop (partial examples shown; full details can be found in Appendix B)

STEP2:

After collecting all the quotes, we categorized them into autonomy types and broke them down into more specific needs or pain points (as shown in *Fig 4.6.3*). **This allowed us to understand the primary needs of design students within this context.**

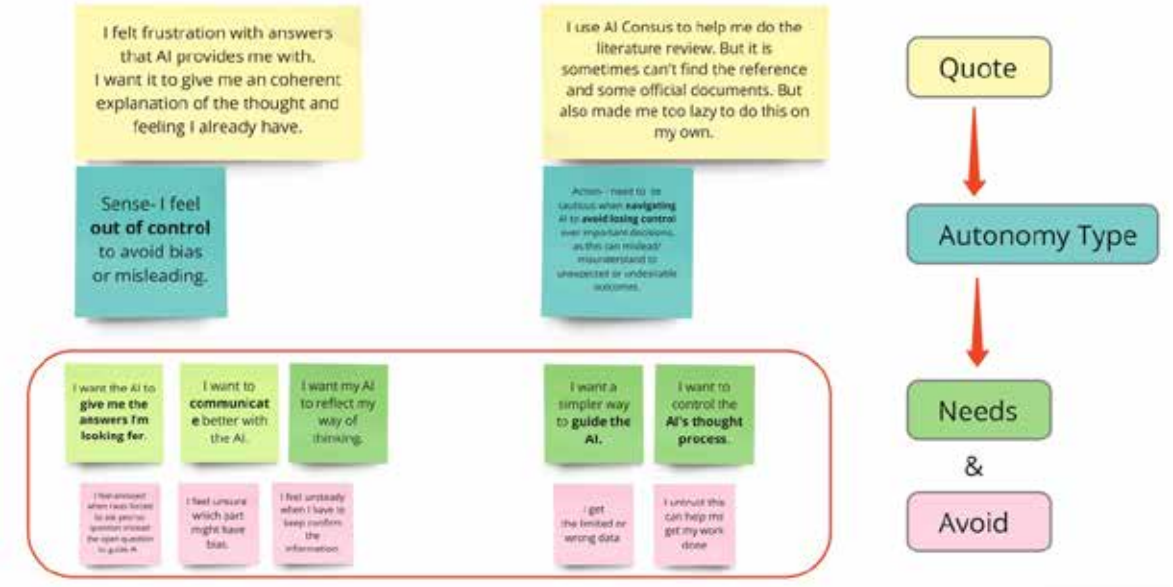


Fig 4.6.3 Quotes break down into needs & pain points

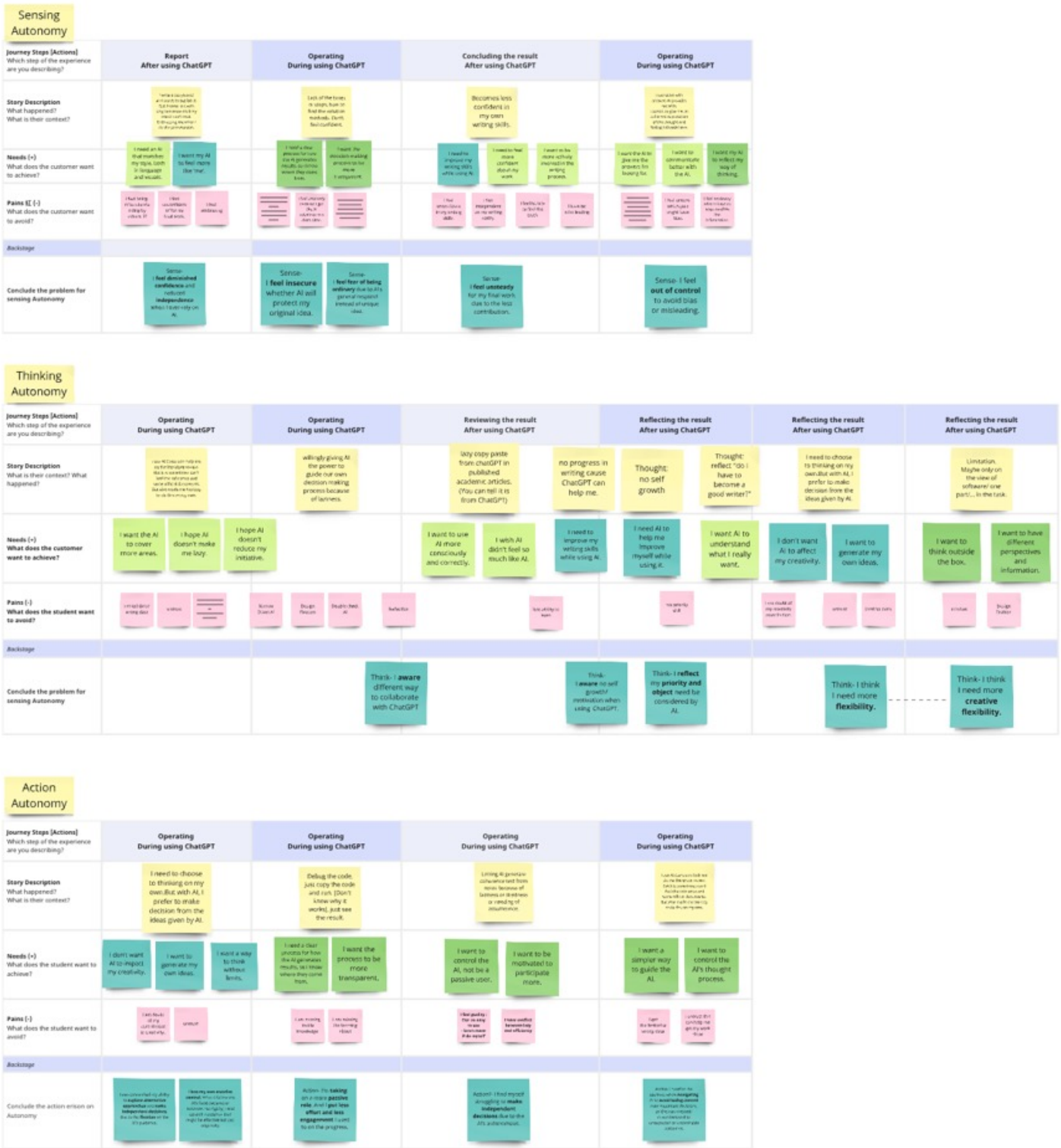


Fig4.6.4 All Quotes analyzed from the workshop broken down into needs and pains (full details can be found in Appendix B)

STEP3:

The analysis categorizes the design needs into three main types, noting that some areas have overlaps while others are more distinct. (Figure 9)

Findings:

1. Primary Design Needs:

- The highest priority is placed on **Thinking Autonomy**, as this area shows the most significant overlaps and recurrent intersections with **Action and Sense** autonomy.
- The dark green areas within the Thinking category indicate common needs and hold significant research value.

2. Secondary Design Requirements:

- The needs identified within the **Action Autonomy** category were mentioned at least twice.
- Although the sample size is limited, these recurring needs can be considered as secondary design requirements to support the overall design

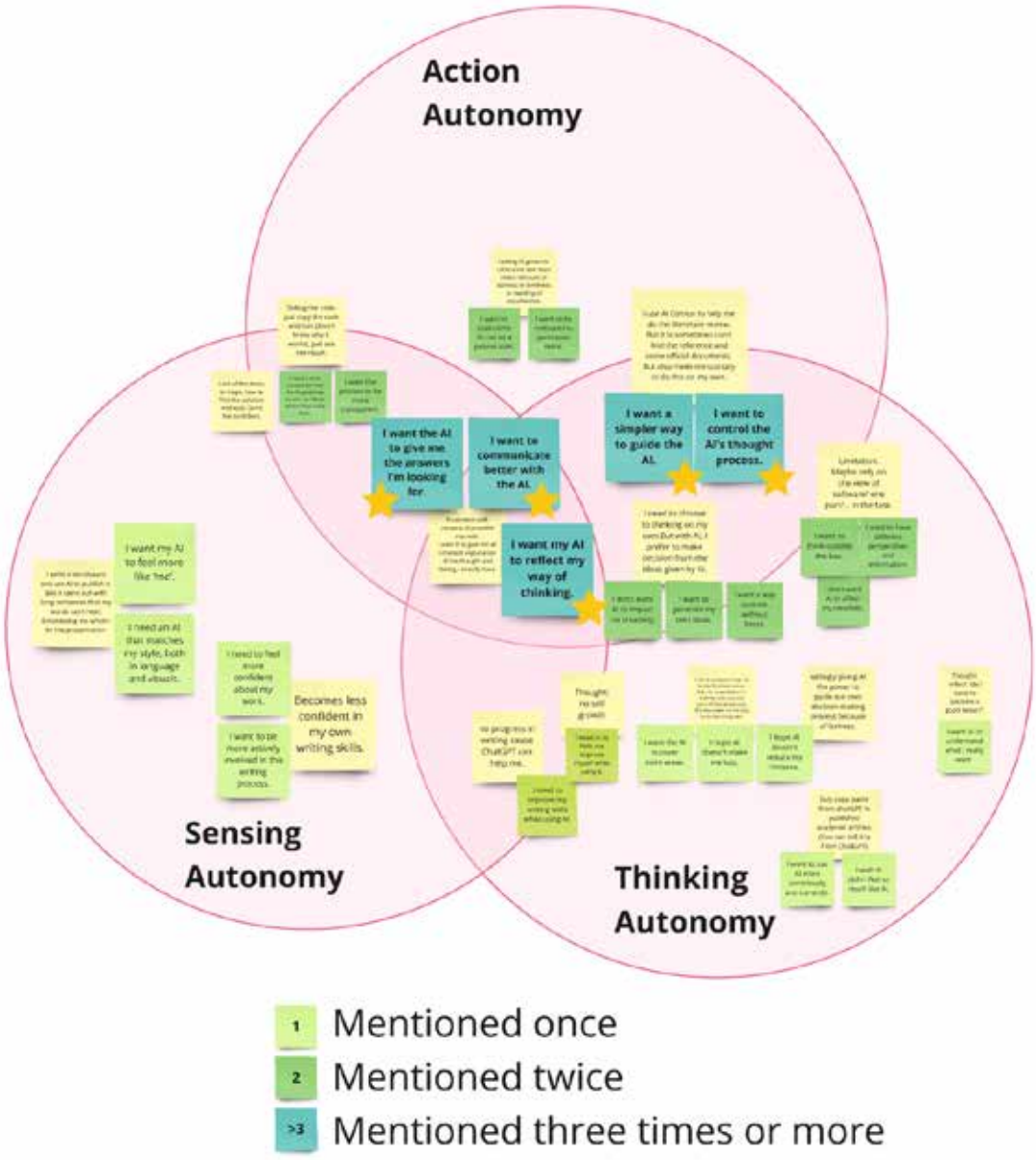


Fig 4.6.5 Three main types and the frequency of mentions

STEP4:

Building on the previous step, we highlighted the dark green Post-its with the highest frequency of mentions, while the medium green Post-its, mentioned more than twice, served as supporting concepts. This helped us consolidate the users' primary needs, which were then used to create the design criteria as a direction for future design (as shown in Figure 10).

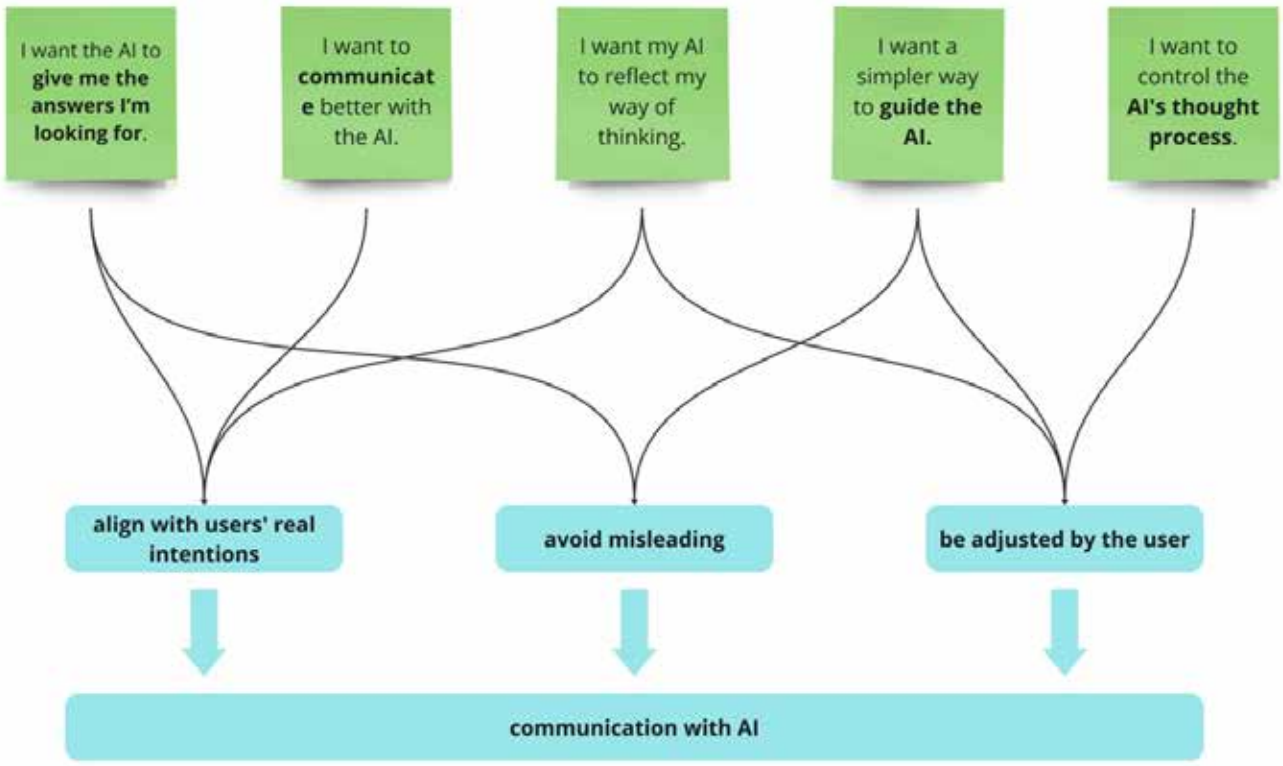


Fig 4.6.6 Translate Needs into design criteria

Finally, we established clear design criteria, which can be consolidated into a comprehensive design main goal and sub-goals. In the next chapter, CH5, we will provide a detailed explanation of the problem statement, design goal, and the three design criteria to facilitate the subsequent design ideation process.

5 Design

5.1 Problem statement

In the rapidly evolving field of human-AI interaction design, large language models (LLMs) like ChatGPT present both potential and limitations as design support tools. Research indicates that while ChatGPT can effectively perform tasks as a designer, user, or product, it also **suffers from issues like information gaps and repetitive responses** (Kocaballi, 2023). ChatGPT has proven useful in various stages of the **design thinking process**, providing human-centered and emotional insights that foster innovative solutions (Fischer et al., 2023). However, despite saving time and **structuring content as a brainstorming tool**, it often generates shallow content, **requiring extensive prompting, cajoling, and manual editing** to achieve desired results (Meron & Araci, 2023). Additionally, while ChatGPT aids design knowledge management, the quality of information retrieval heavily **relies on precise prompts** (Hu et al., 2023).

Our previous workshop revealed the insight that users seek more accurate ways to express their needs, expecting the AI to align closely with their true intentions. Current AI systems often **create communication gaps due to inconsistent information and vague responses**, making it challenging for users to have a satisfactory experience. Users also desire **greater control to adjust AI outputs** to better meet personal needs, but **current systems lack the flexibility needed for practical and personalized interaction. [Gap]** In effect, these limitations undermine the collaborative potential of human-AI interaction, restricting AI's role in design. Therefore, there is a need for a more human-centered and flexible AI communication system that **aligns with user intentions, avoids misleading communication, and offers more adjustable options** to enhance collaboration accuracy and satisfaction. Such a system would support effective knowledge management and emotional engagement, enabling a more efficient and reliable design process.

5.2 Design Directions

Approach

Affordance Concept in Human-AI Interaction

Shift from "humans adapting to AI" to "AI adapting to humans"

In current human-AI interactions, people often need to continually adjust and adapt their prompts to elicit responses from ChatGPT that better meet their needs. According to Don Norman's affordance theory, we can rethink this interaction, designing ChatGPT to more naturally adapt to human expressions, enabling a smoother communication experience.

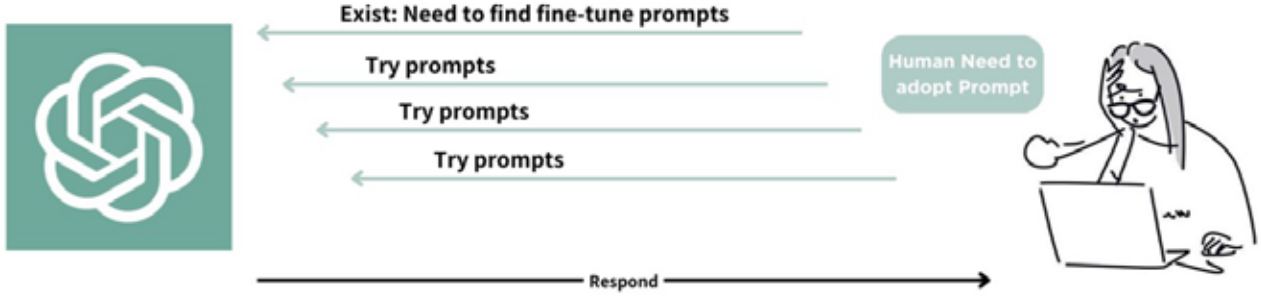


Fig 5.2.1 The Interaction with ChatGPT nowadays

Shift from "humans adapting to AI" to "AI adapting to humans"

- Human affordances: Humans have natural ways of expressing intent, delicate emotions, and personalized needs, which AI systems may not inherently understand.
- ChatGPT affordances: ChatGPT's responses are currently more fixed, requiring users to fine-tune prompts to achieve ideal outcomes. This interaction style often results in information gaps and repetitive adjustments.
- Ideal Design: Future versions of ChatGPT should develop **affordances that adapt to human expressions** more flexibly, reducing reliance on precise prompts. This shift would transform interactions into a model where the **AI adapts to human expressions**.

Fig 5.2.2 The alternative way to interact with ChatGPT

Main Design Goal

Promote a user experience that enables humans to communicate their ideas more effectively to an ideation AI system, ensuring a productive and intuitive human-AI collaboration.

Sub Goals

1. Align with Users' Real Intentions: Ensure that the AI system accurately understands and aligns with the authentic intentions and goals of the users.

2. Avoid Misleading Communication: Design interactions in a way that prevents misinterpretation or misleading outputs from the AI, promoting clarity and reliability in the AI's responses.

3. Allow User Adjustments: Provide users with the ability to refine and adjust AI outputs, ensuring

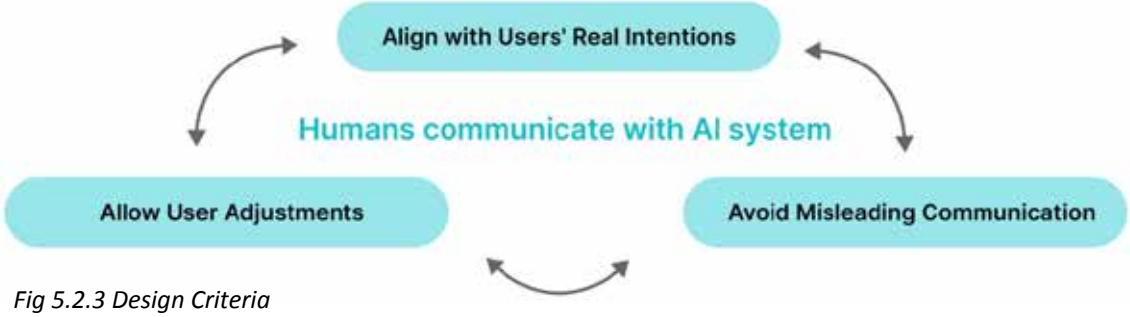


Fig 5.2.3 Design Criteria

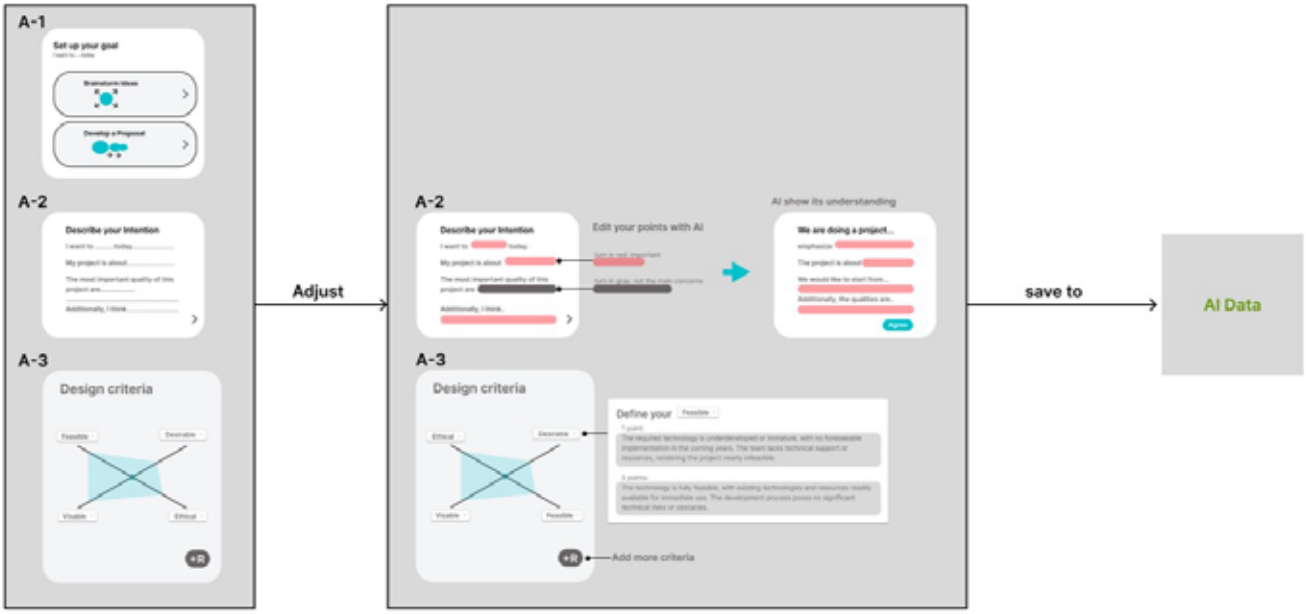
5.3 Process of conceptualization: 2 Round

First Version

A. Align with Users' Real Intentions

Visualizing AI Understanding: Enhancing Communication and Interaction through Transparent Interpretation

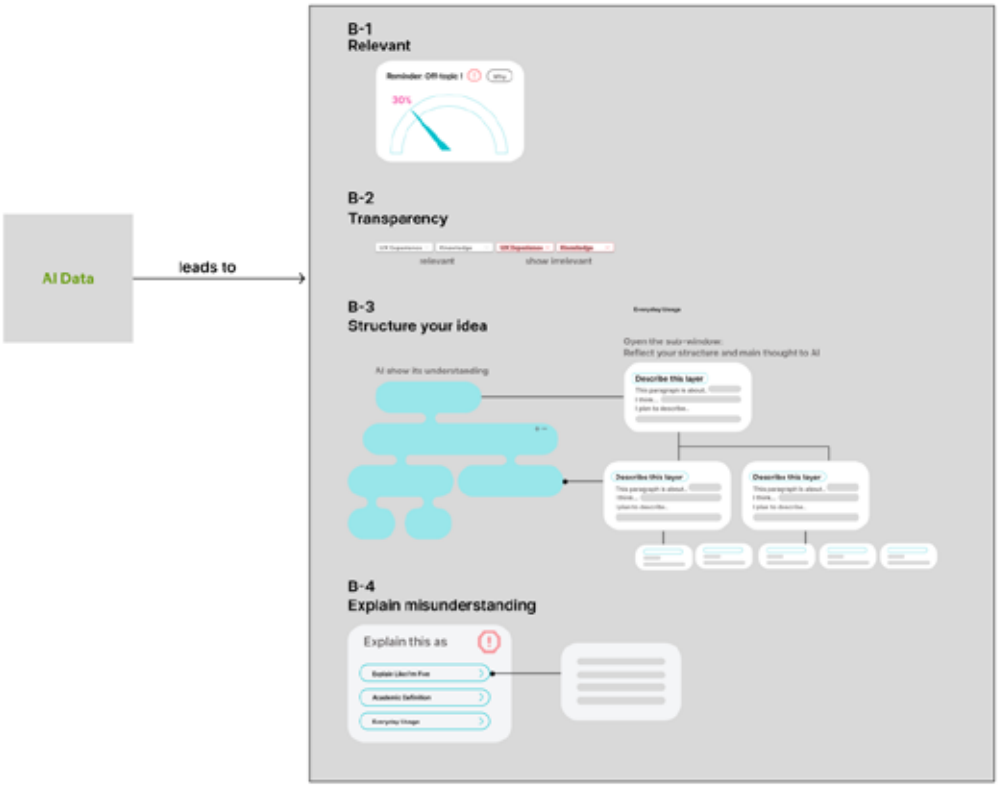
Imagine a scenario where we can visualize how an AI system interprets our priorities, intentions, or the quality of our inputs. These features would allow us to reflect on our project by first clarifying and expressing our needs to the AI and then confirming that the AI has understood them correctly.



Avoid Misleading Communication

Maintaining Focus in AI Conversations: Redirecting Off-Topic Dialogue for Effective Interaction

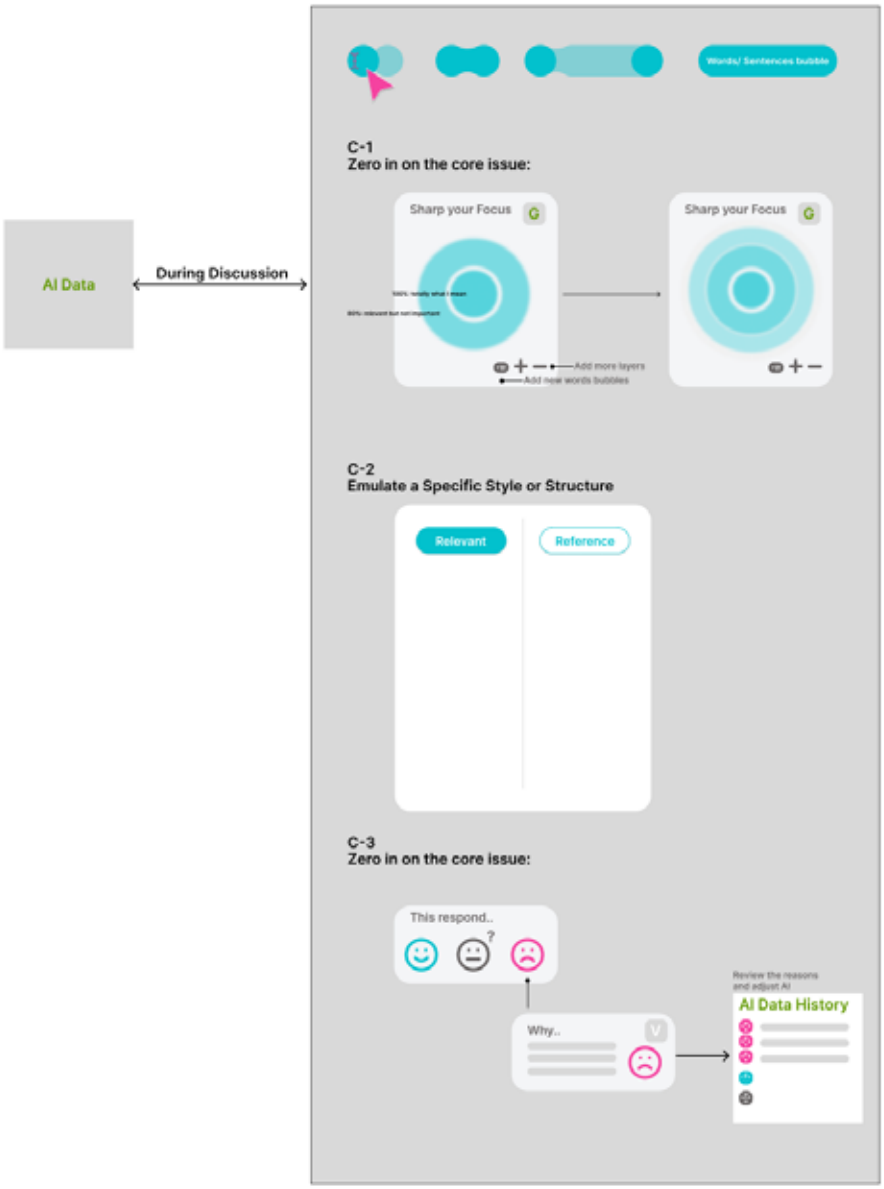
Imagine a scenario where you're discussing a specific topic with an AI, but the AI starts giving irrelevant suggestions, taking the conversation off course and causing confusion.
To avoid this, the AI could recognize when the conversation shifts away from the main topic and remind you to return to the original discussion. Similarly, the user can also notice the off-topic drift and actively redirect the conversation back on track, ensuring a focused and productive exchange.

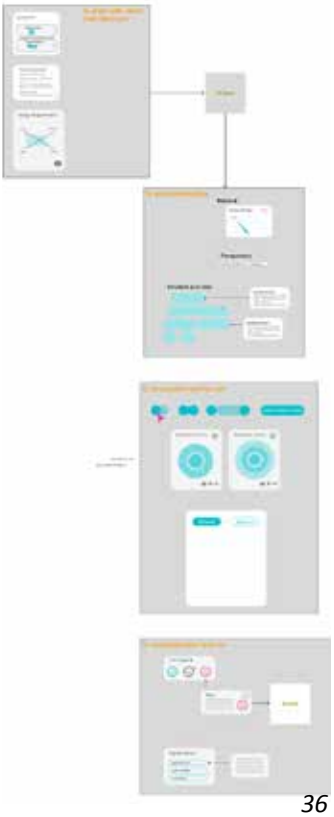


Allow User Adjustments

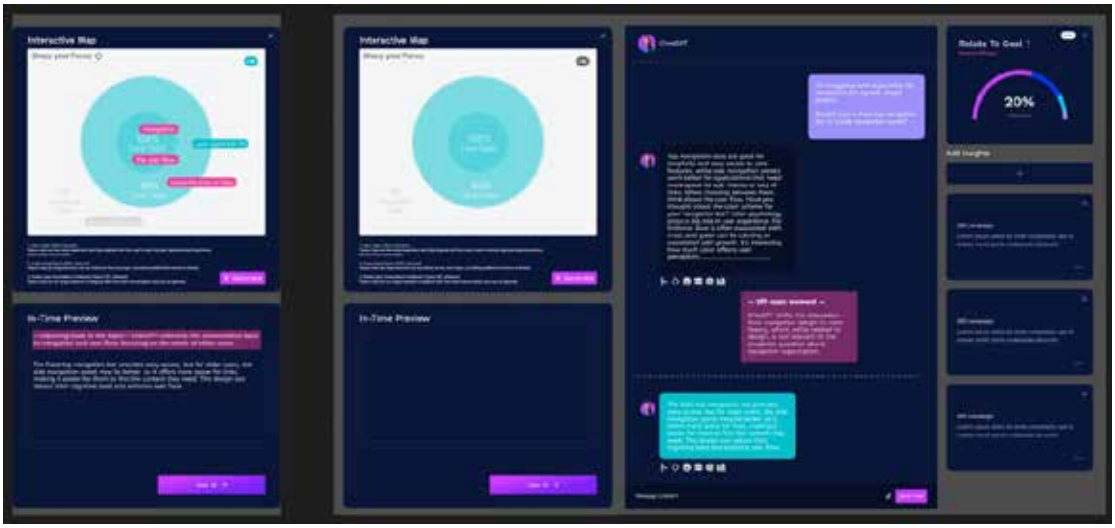
Allowing User Adjustments for Focused AI Interaction

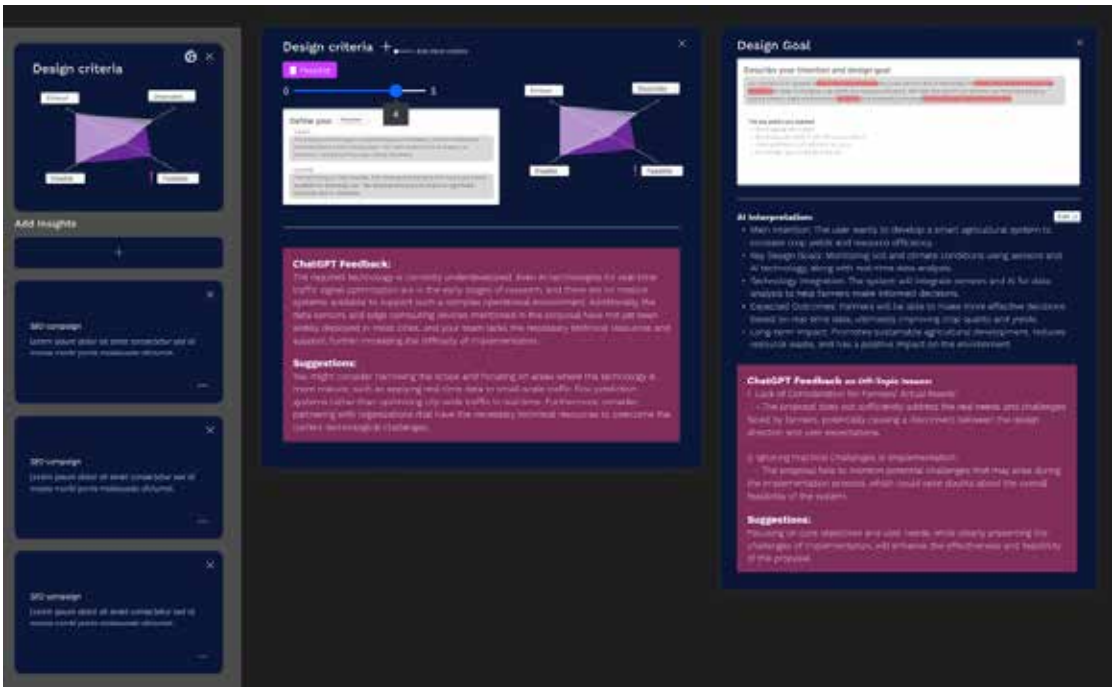
Imagine a scenario where you're discussing a topic with an AI, but its responses don't fully align with your priorities. To improve the interaction, the AI allows you to adjust its focus, such as prioritizing certain aspects of the discussion or shifting the level of detail.
With these adjustments, the AI can adapt in real-time, ensuring the conversation stays relevant and tailored to your needs.

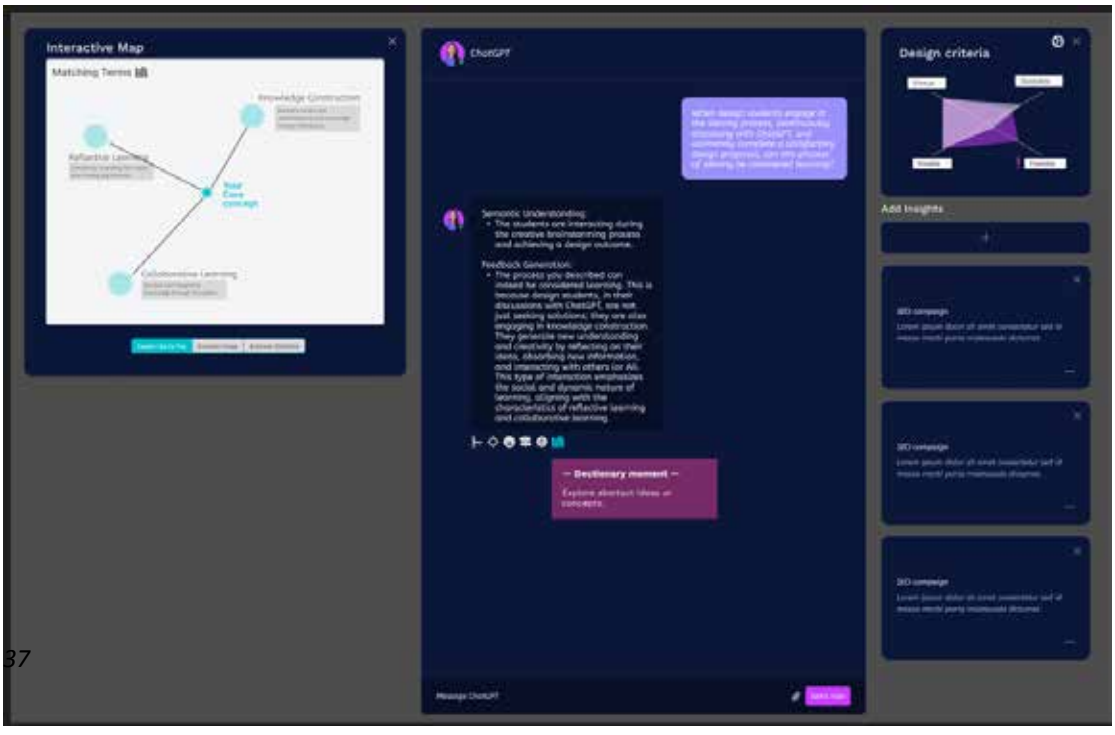


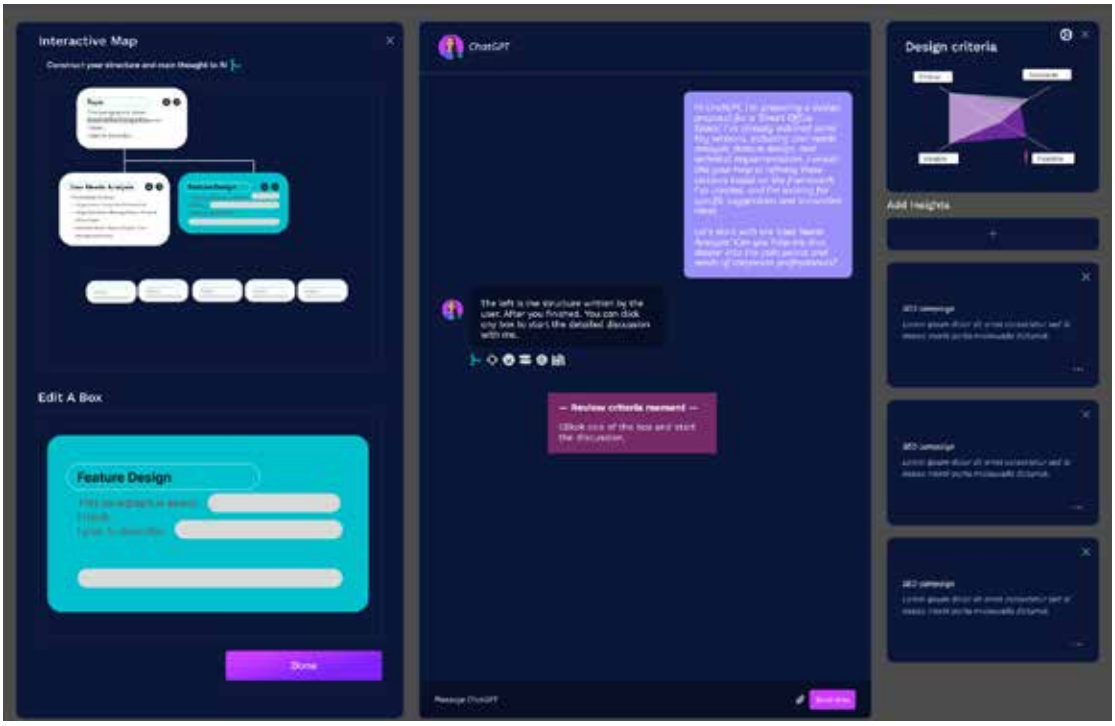


Second Version









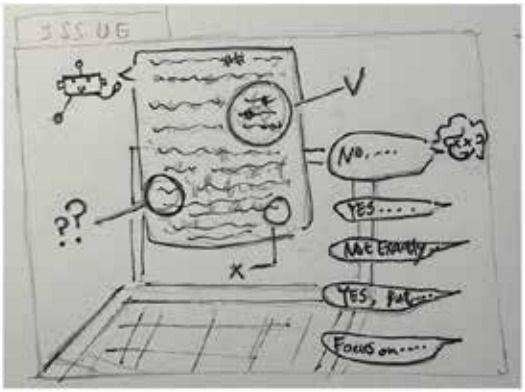
Quick Take away (What Change)

function bar (left only top one)
the order for operating
adds bookmark in Dictionary
the Note card color and status
Adds the cards functions (select/ trace back)
Big card review status
Change icons
Make the style in wireframe (little color)

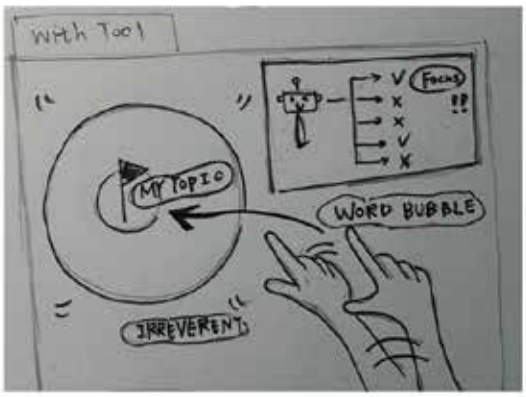
5.4 Final Design Concept Proposal

Final Version

Function1: Off- Topic
Criteria: Align with users' real intentions, Be adjusted by the user

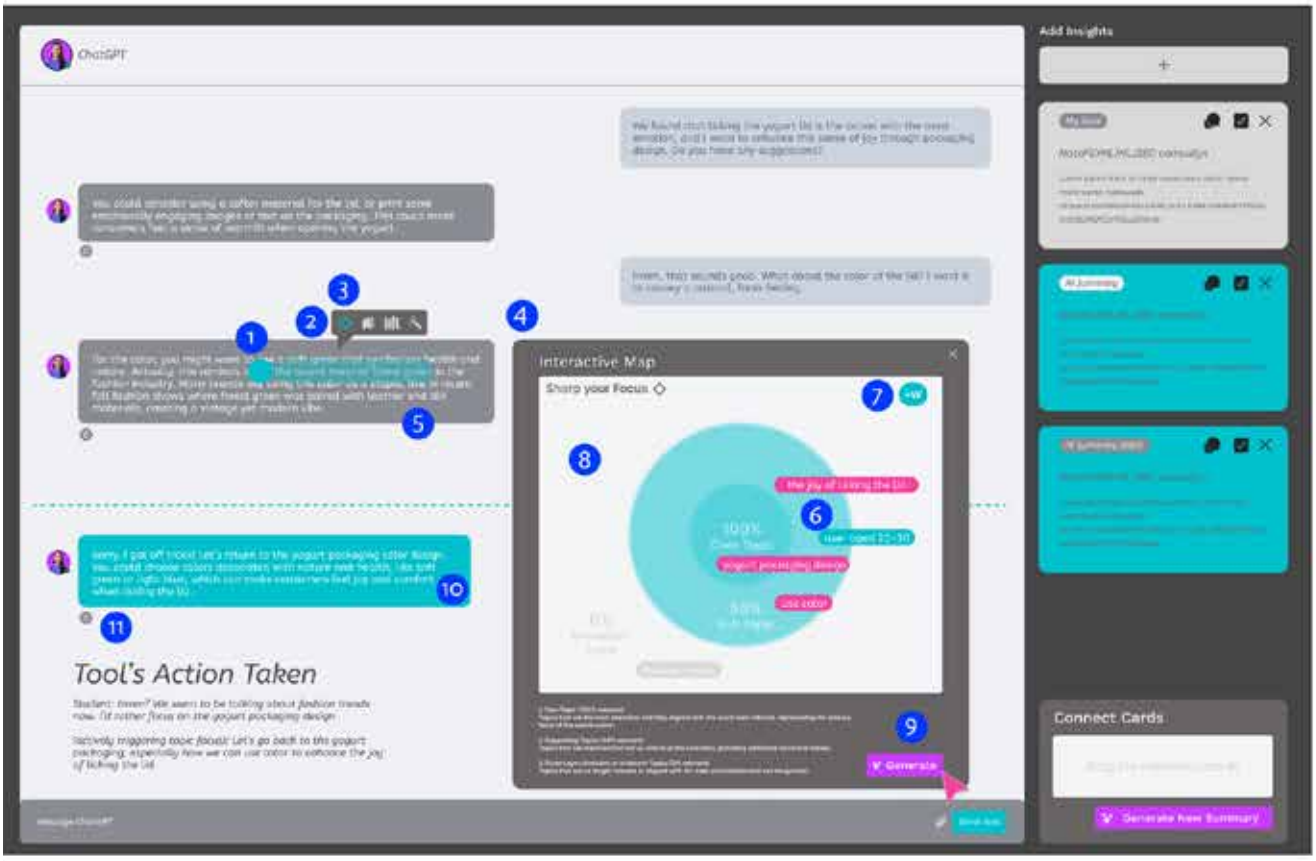


When ChatGPT provides a lot of irrelevant information or goes off-topic, I want it to stay focused and give concise, accurate responses.



What if

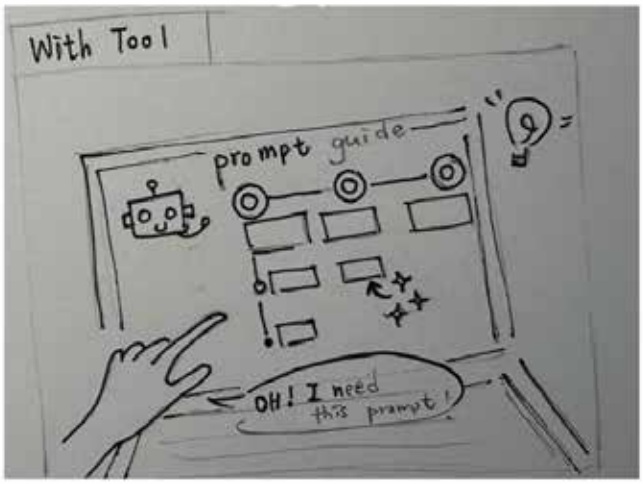
we could communicate our key points to AI more intuitively, so it understands exactly what we're focusing on and responds accordingly?



1. Select key text
2. A floating toolbox will appear in the upper right corner
3. Click on ⊙
4. An Interactive Map window will pop up
5. Return to the conversation and drag the keyword bubbles onto the Map
6. Repeat step 5 to add more topic keyword bubbles
7. (You can press +W to add topic bubbles manually)
8. Arrange the bubbles' positions on the Interactive Map by moving them around
9. Press the Generate button
10. ChatGPT will respond based on the focus topics on the Map
11. If you're not satisfied, you can Generate again or press the ⟳ button

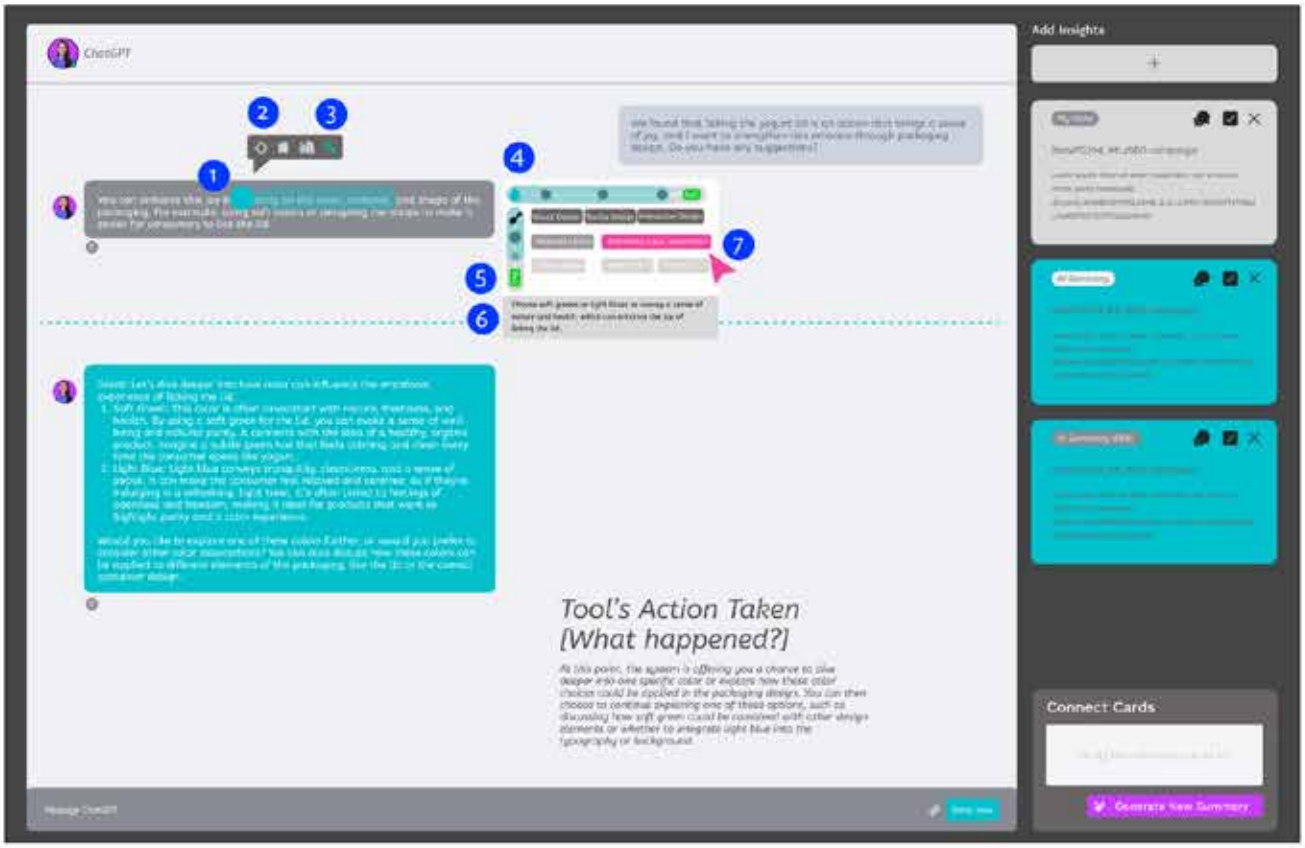
Function2: Prompt Guide
Criteria: Align with users' real intentions

When I have many ideas but don't know how to give clear instructions, I want the AI to understand my intentions and provide focused, relevant responses.



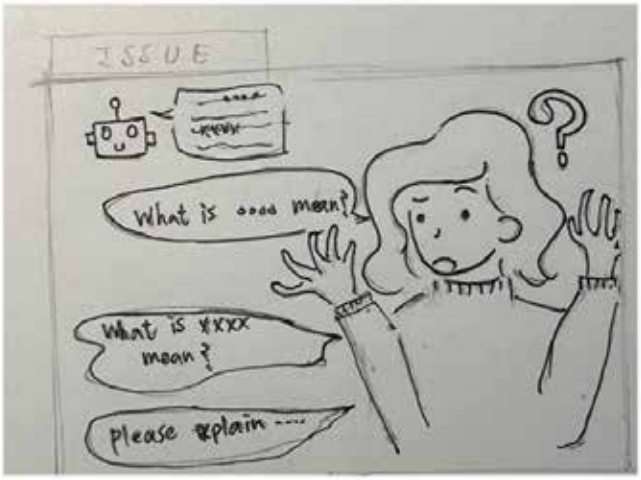
What if

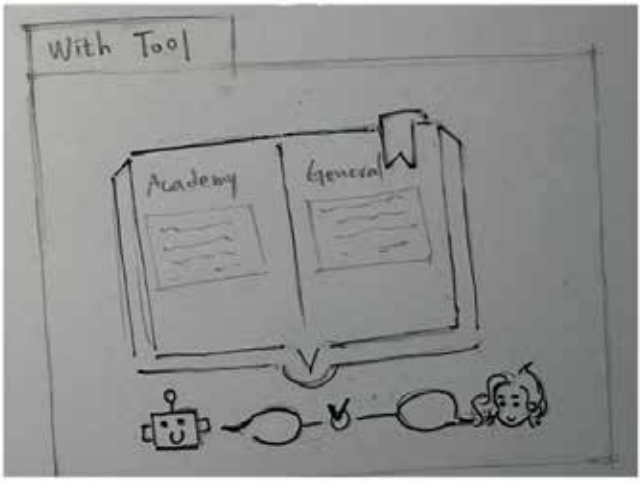
we could communicate our intentions to AI more clearly, so it understands and responds accurately?



1. Select key text
2. A floating toolbox will appear in the upper right corner
3. Click on
4. An Prompt Guide window will pop up
5. Slide the depth lever to explore deeper keywords
6. Relevant prompts will display at the bottom of the window
7. Select the desired keyword, and ChatGPT will respond based on it

Function3: Dictionary
Criteria: Avoid misleading communication

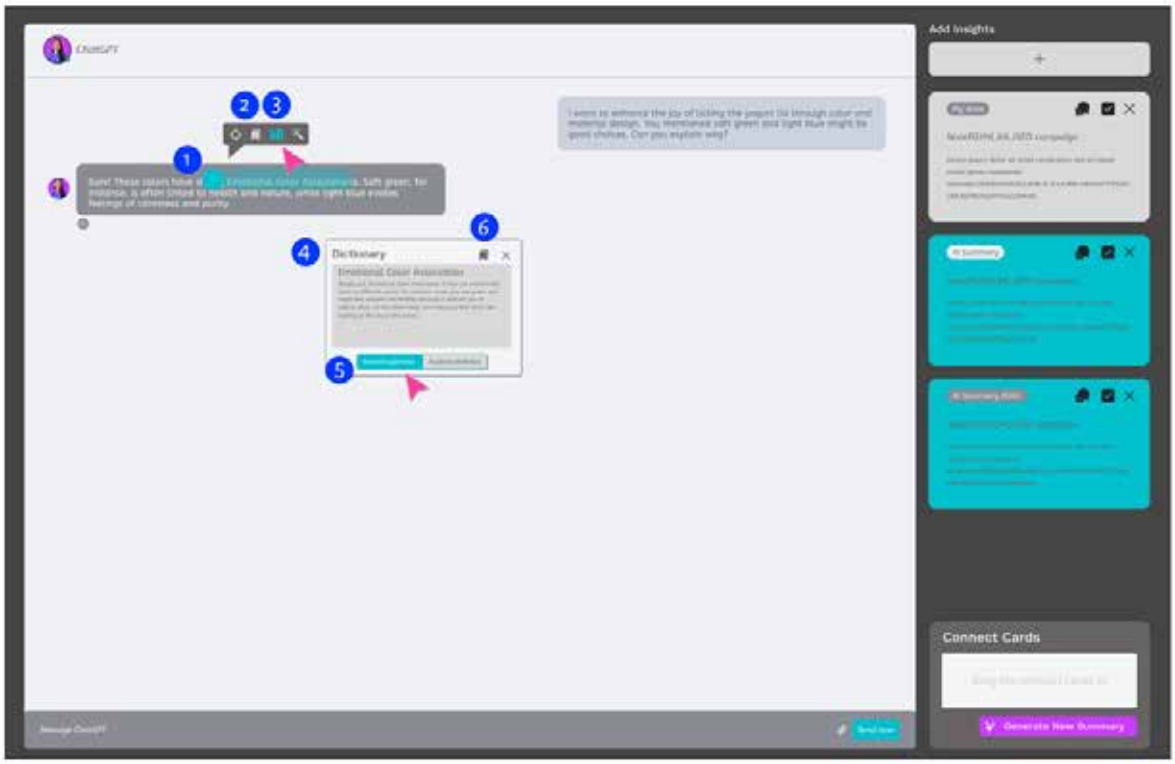




When the AI uses obscure words or complex concepts, I want it to automatically provide definitions or explanations, so I can avoid misunderstandings and keep the conversation clear."

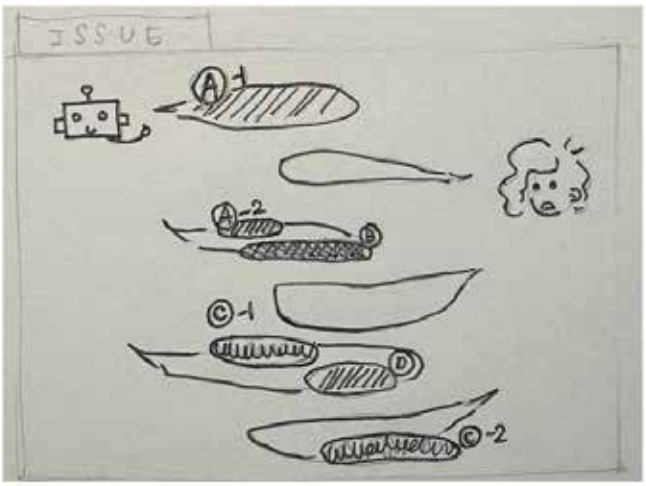
What if

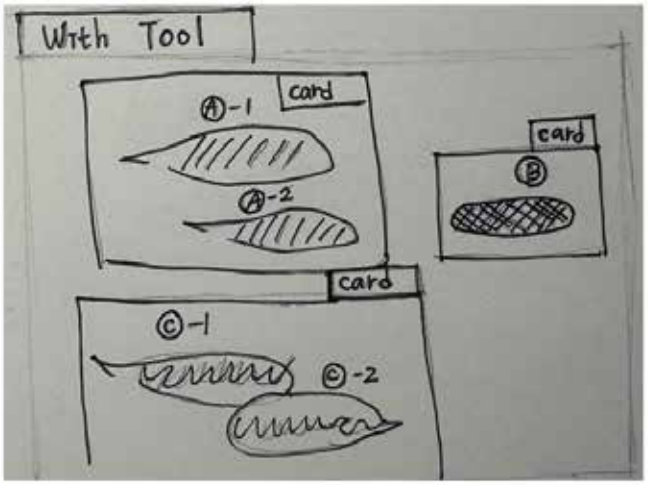
we could correctly understand the language AI is trying to express, avoiding any misunderstandings in the process?"



1. Select key text
2. A floating toolbox will appear in the upper right corner
3. Click on
4. An Dictionary window will pop up
5. Select General explanation or Academic definition button
6. (Add to note if needed)

Function4: Card Note
Criteria: Align with users' real intentions, Be adjusted by the user

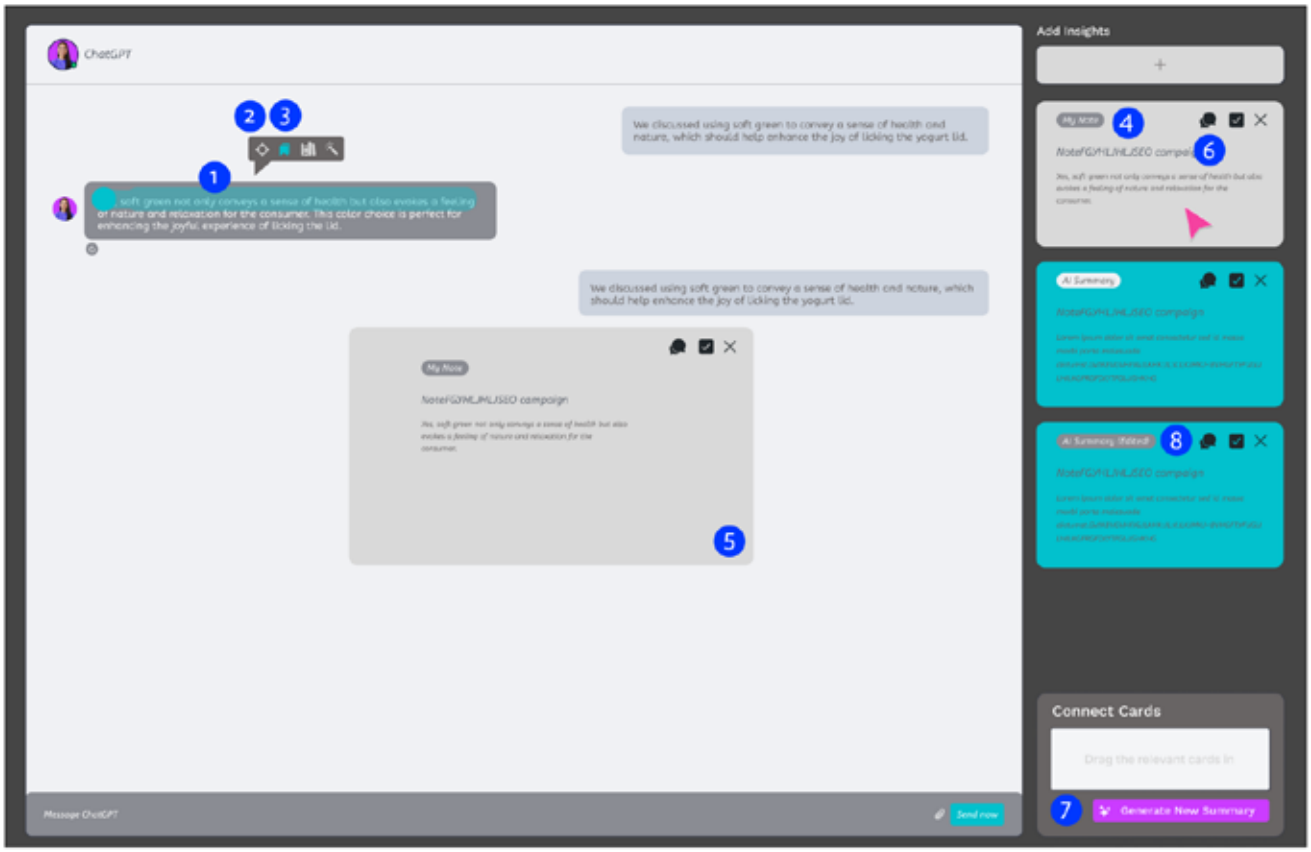




When having continuous discussions with AI, I often need to take notes while thinking. Sometimes, I miss recording certain points and have to scroll back to review or re-explain things we've already discussed to the AI. Additionally, the topics may be scattered across various conversations, or the order might be inconsistent, making it difficult to revisit the overall context.

What if

we could trace back and logically organize our thought process, ensuring clarity and better understanding of how we arrived at a conclusion?



1. Select a specific paragraph of text
2. A floating toolbox will appear in the upper right corner
3. Click on
4. The selected paragraph will become a note, recorded on a card on the right labeled "My Note."
5. Click on a card in the right sidebar to open and view its content.
6. Press a button to quickly revert to previous discussion threads.
7. If multiple cards in the right sidebar are selected, drag them to the lower right to merge them into an "AI Summary."
8. Click on the "AI Summary" card to further edit it with your own input.

6

Evaluation

CH6 Evaluation

6.1 Evaluation plan

Time: One-on-one session, 45 minutes
Interview Subjects: IDE design students aged 25–30 (26F, 26M, 29F, 28F)

- Warm-up Interview:

Reflection 1: ChatGPT usage habits and communication style (represented in a graphical form, if possible).
Reflection 2: Challenges faced in usage
[Note: Avoid mentioning design criteria (real intention, user adjustments, misleading) here to prevent excessive prompting].

- Part 1: Introduction of 4 Features

For each feature, provide (1) usage scenario, (2) "what if" proposal, (3) feature introduction and basic operation, and (4) collect feedback, including:
a. Inspiration: Relation to and comparison with past usage experiences (first understand how users previously approached similar tasks).
b. Helpfulness: Whether the feature meets the typical needs of design students.
c. Usability: The intuitive recognizability of the feature's interface and interaction.
d. Design Improvement: Good aspects of the design or areas for further adjustment.
e. Favorite Feature: Which feature is liked the most, and why?

- Part 2: Use the **"Think Out Loud"** method to operate two features (Off-Topic and Notebook). Participants will freely interact with the features based on their memory from Part 1 - Introduction Brief, while verbalizing their actions and thoughts. The interviewer will ask **"why"** to understand the reasoning behind actions and gather feedback, including operational smoothness, interaction feedback, whether they were correctly guided and operated as intended, feature recognizability, icon recognizability, and directions for future iteration or improvement.

6.2 Results and Discussion

In this design thesis chapter, the **Framework Analysis method** is applied to evaluate the alignment of four key features with the design goals and criteria established from previous research. This analysis aims to determine **whether these features fulfill the intended purpose of promoting effective and intuitive human-AI collaboration.**

Based on the Chapter 5.2 research, our design goal focuses on enabling a user experience that facilitates seamless communication between users and the AI system. Three sub-goals are established to guide this:
1. Align with users' real intentions: Ensuring the AI accurately interprets and supports the user's authentic goals.
2. Avoid misleading communication: Promoting clarity and minimizing miscommunication between users and the AI.
3. Allow user adjustments: Allowing flexibility for users to refine and adjust outputs based on their needs.

Four features (F1: Off Topic, F2: Prompt Guide, F3: Dictionary, F4: Card Note) were developed and tested to address these goals. Every feature might include multiple criteria; the graph below shows the link.

Function1: Off- Topic
Criteria: Align with users' real intentions, Be adjusted by the user

Align with users' real intentions

(V) Helping users focus on key conversation topics

The Off-Topic feature enables users to quickly focus on core topics and reduce distractions caused by irrelevant content. Participant A mentioned, *"This feature is novel, and I can understand it very well! I often encounter situations like this. However, I don't consider this off-topic but rather the AI's own thoughts."*

She further explained that by emphasizing or reducing specific content, she could generate outputs that better fit her needs. This indicates that users can naturally express their intentions by adjusting the generated content through the feature, which aligns with the design criterion Align with users' real intentions.

(V) Enhanced engagement in conversations

Participant C reflected on her past experience when AI-generated content deviated from the intended topic. She had to rely on repeated trial-and-error inputs or reminders to guide the AI back to the main topic, which was often time-consuming and frustrating, even causing her to give up and seek human assistance. This reveals the limitations of traditional interaction methods in maintaining topic focus. With the Off-Topic feature, users can intuitively adjust generated content without redundant inputs, significantly improving conversational engagement and efficiency.

(V) Smoother interactions

Participant D shared that, in the past, she had to open a new window to avoid contaminating the existing conversation with unrelated content. However, this method disrupted the continuity of discussions and often led to distraction. She stated, *"This innovative design provides a more intuitive usage experience, reducing unnecessary actions such as constantly starting new conversations."* By allowing users to adjust and filter content seamlessly within the same window, the Off-Topic feature reduces cognitive load and unnecessary steps, thereby maintaining a smooth interaction flow.

(V) Improving clarity in expressing user intentions

The feature also helps users clarify their actual needs and convey them more naturally to the AI. Participant B stated, *"Every time I asked ChatGPT a question before, I had to explain my background information in great detail—for instance, what design goal I wanted to achieve or which audience I was targeting—and I would make sure repeatedly that it understood what I was talking about before continuing the discussion."*

With the Off-Topic feature, she explained, *"Now I can more clearly identify unnecessary content and directly influence the next generated response."*

This demonstrates the feature's strength in enabling users to articulate their needs effectively, significantly reducing the burden of trial-and-error inputs.

Allow User Adjustments

(V) Flexibly adjusting the conversation direction

Users generally found the Off-Topic feature to excel in content adjustment. Participant B noted, *"The drag-and-drop functionality, where I can turn text into bubbles and place them in the target area, is very intuitive. It makes it easier for me to focus on key points."* This drag-and-drop mechanism allows users to prioritize and manage generated content according to their needs, offering intuitive and flexible interaction that reflects a user-centered design approach.

(V) Reducing distractions from irrelevant content

The feature helps users ignore unnecessary information by allowing them to label irrelevant content, thereby reducing cognitive load. Participant C commented, *"I can quickly filter out unrelated content and refocus on the main topic."* This proves particularly effective in scenarios where users need to locate core information from a large volume of generated content.

(X) The need for more efficient handling of large volumes of content

While the feature demonstrates exceptional intuitiveness and flexibility, manual adjustments, such as dragging and dropping, may become inefficient in high-information scenarios. Participant C suggested, *"It would be great if the feature could leverage AI to automatically detect and prioritize key points in the generated content, thereby improving operational efficiency."*

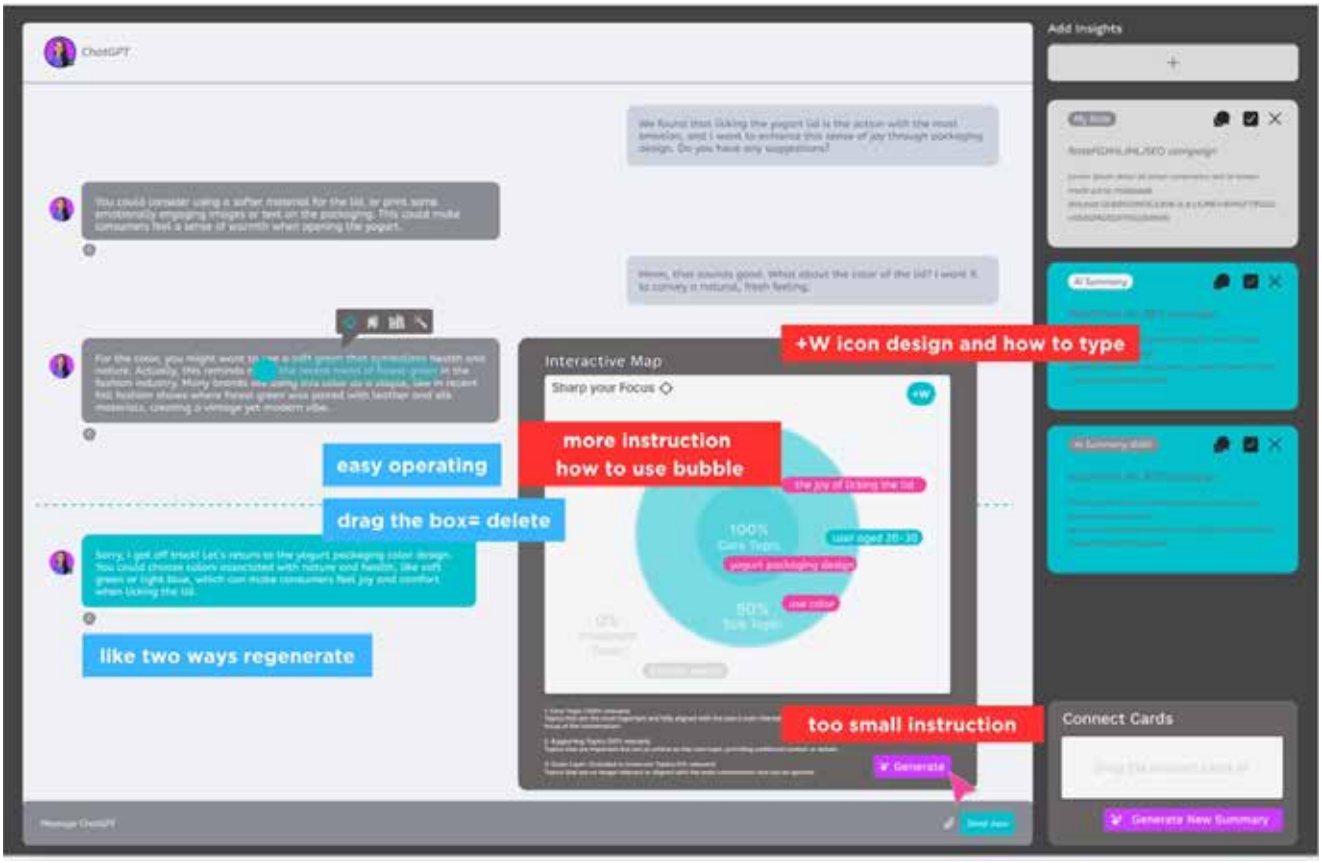
Conclusion and Suggestions for Improvement

The Off-Topic feature excels in supporting users to express their real intentions, adjust content flexibly, and maintain smooth conversations. Most users appreciated the feature's innovation and intuitiveness, particularly its simplicity and effectiveness in helping them focus on key topics during conversations. However, there is room for improvement in handling high-information scenarios. Some users suggested introducing batch selection or intelligent filtering tools to enhance efficiency. Additionally, enhancing visual feedback, such as dynamic animations or color changes, could further improve the user experience.

UI & UX Feedback

In terms of interaction, 3 out of 4 users found the operation very smooth. They were able to intuitively understand each step and the next action, arranging the "text bubbles" into the interaction window one by one to complete the topic redirecting action. Only one user, C, mentioned feeling unsure of what to do next when the interaction window popped up. She suggested making the guidance for the bubble function clearer, such as by adding an introductory animation to explain how to select and remove bubbles. User D further suggested that the text selected by the bubbles could appear more three-dimensional visually, such as with a floating animation or shadow effect for a more pronounced look. Additionally, User B noted that the explanatory text at the bottom was too small and could be made clearer by displaying an additional explanatory window when the mouse hovers over it. Lastly, in the "add emphasis" action using “+W”, the Think Out Loud process revealed a need for a more clearly defined input box and a suitable representative icon.

Overall, the four users had a positive attitude toward the new interaction model, finding it interesting and innovative. They appreciated the intuitive visual operations and interactions, considering them quite inspiring. They enjoyed being able to rearrange the "text bubbles" in the interactive map and regenerate responses. They also highlighted the importance of retaining the original "regenerate" button in the reply box, with User A commenting, "After all, it’s a different approach, and I like having options based on the current situation."



Function2: Prompt Guide
Criteria: Align with users' real intentions

Align with users' real intentions

(V) Supporting brainstorming to inspire ideas

The Prompt Guide feature demonstrates its capability to assist users in brainstorming, especially when they lack specific questions or goals. Participant A mentioned, *"Sometimes I don't even know what to brainstorm about, and this feature helps me wander around and explore ideas."* This indicates that the feature provides a structured way for users to initiate creative thinking. By allowing users to explore loosely connected topics, the feature aligns with the design goal of enabling users to uncover their real intentions through wandering and inspiration.

At the same time, most participants highlighted the importance of having specific goals when using AI for brainstorming. Participant A also noted, *"When I want ChatGPT to brainstorm something, I at least know what to ask."* This reflects the dual nature of the feature: while it effectively supports unstructured exploration, it remains dependent on users having an initial framework or starting point to achieve optimal results.

(O) Reducing trial-and-error inputs

The feature also partially supports users in minimizing iterative inputs by providing a visualized framework for clearer and more direct interactions. Participant D commented, *"If I don't like the result, I just naturally adjust and tell it again."* This reflects her reliance on trial-and-error inputs to refine AI responses. However, D also pointed out, *"It depends on the situation. Sometimes I just tell it to modify something, or I rewrite the entire prompt—it depends on which feels easier."* While this feedback indicates habitual reliance on iterative methods, it also suggests that the feature might feel unnecessary or cumbersome for users accustomed to refining their prompts organically.

Overall, while the Prompt Guide feature offers value in visualizing potential inputs, its effectiveness in reducing trial-and-error remains context-dependent. The feature seems better suited to short-term, exploratory tasks, such as brainstorming, rather than long-term, cumulative discussions.

(X) Lack of specificity in prompt suggestions

Some participants expressed dissatisfaction with the generalized nature of the prompt suggestions, which occasionally led to irrelevant or imprecise outputs. Participant C shared, *"If you ask it to tweak something into another style, it might deviate from the original topic."* This demonstrates the risk of broad suggestions causing the AI to stray from the user's original intention, making it harder to maintain focus. Additionally, Participant B questioned the necessity of the feature, stating, *"Is adjusting your prompt based on its response really that difficult? Do we need something to assist with this?"* This skepticism highlights a potential mismatch between the feature's design intent and the expectations of more experienced users, who prefer precision and direct control over their interactions.

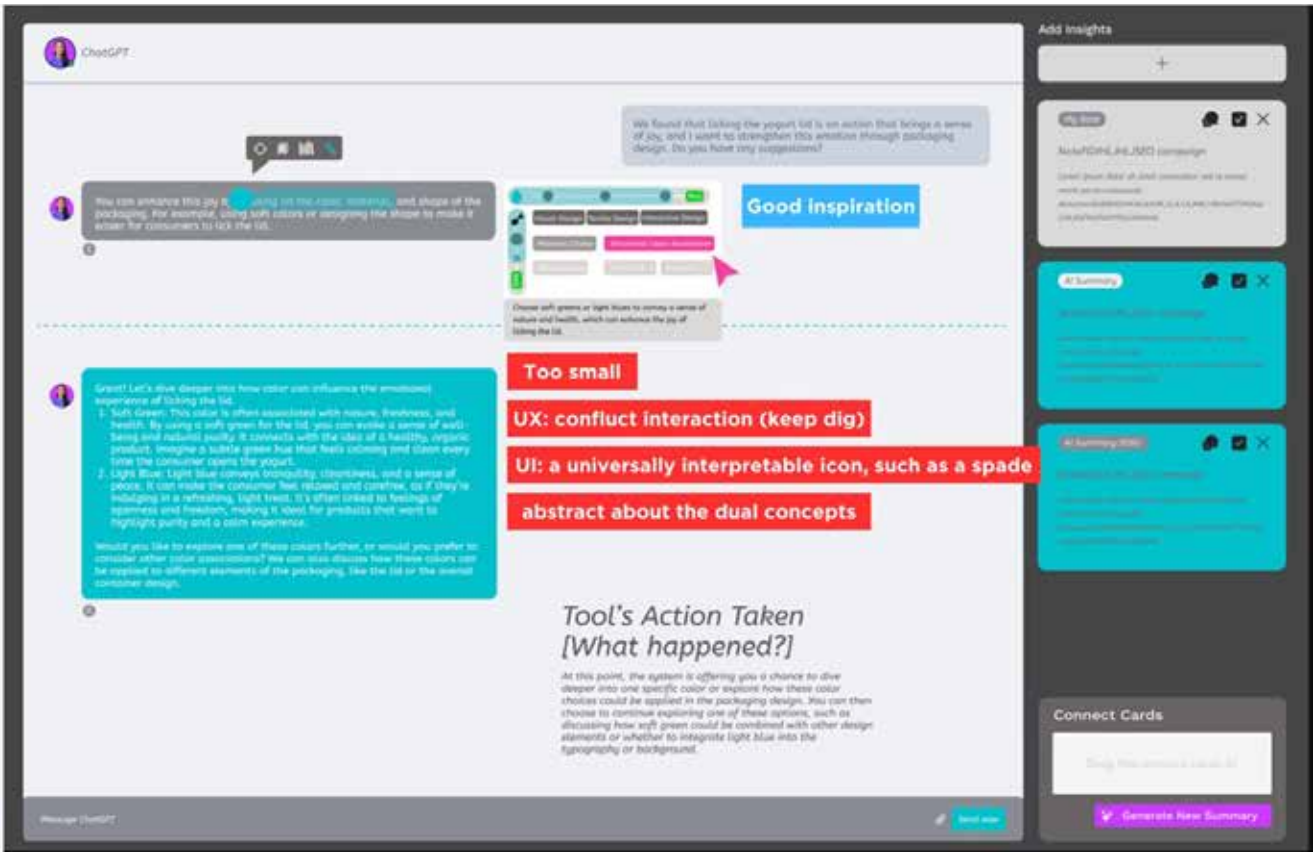
These observations suggest that while the feature may benefit users who struggle with prompt framing, its generalized nature could hinder its effectiveness for tasks requiring high specificity or precise direction.

Discussion and Suggestions

The Prompt Guide feature provides valuable support in early-stage brainstorming by guiding users through wandering and unstructured exploration. It enables users to uncover ideas when they lack a clear starting point, aligning with the design goal of helping users articulate their real intentions. However, its utility is limited in longer, cumulative discussions where users rely on trial-and-error prompts to refine interactions. Additionally, the generalized nature of prompt suggestions may not meet the needs of experienced users who require more tailored and precise guidance. Future iterations of the feature could address these limitations by offering adjustable levels of prompt specificity and optimizing for cumulative, long-term discussions.

UI & UX Feedback

Users A and D could connect with the usage context and pointed out some UI/UX design issues. They felt that the interactive dialogue box was too small, needing more space to operate and explore ideas. Additionally, they suggested that the breadth slider might be unnecessary to simplify the overall interface. User D mentioned, "Maybe just a shovel icon would suffice, as I think it already encompasses the meanings of breadth and depth." In short, the dual sliders for horizontal and vertical adjustments were overly complex and hard to understand in the visual interface and practical use, suggesting a potential area for improvement in the future.



Function3: Dictionary
Criteria: Avoid misleading communication

Avoid misleading communication

(V) Reducing misunderstandings caused by technical terms

The Dictionary feature effectively reduces misunderstandings arising from unfamiliar technical terms. Participants noted that it helped clarify ambiguous or complex terminology, ensuring smoother communication. Participant A remarked, *"This quick function is very straightforward and easy to use, and I can access it whenever I need."* Similarly, Participant B emphasized its value in resolving uncertainties, stating, *"I feel that ChatGPT often forgets the abbreviations or specialized terms we've discussed earlier. With this feature, I don't have to guess repeatedly or double-check if my understanding matches ChatGPT's interpretation."*

All participants (4/4) expressed satisfaction with the Dictionary feature, highlighting its ability to improve clarity and facilitate the recognition and exchange of valuable information. Furthermore, they mentioned that it helped them articulate their thoughts more effectively, enabling more productive conversations. This aligns with the design goal of avoiding misleading communication by minimizing ambiguity and reducing cognitive load.

(X) Limited contextual adaptability

Despite its effectiveness, some participants (2/4) raised concerns about the feature's limited contextual adaptability. They pointed out that the Dictionary primarily provides generic definitions of terms, which may not always align with the specific context of ongoing discussions. Participant C shared, *"After chatting with GPT for several days, I often feel that certain phrases or terms in the conversation hold unique value. I would prefer to define them in context and save these definitions for future reference."* This feedback highlights a potential gap between the feature's current capabilities and users' expectations. While the Dictionary helps with general term clarification, participants suggested that an "explanation" feature—providing context-specific definitions based on the ongoing discussion—would be more beneficial. This would enhance the functionality by aligning definitions with the immediate conversational context and enabling users to bookmark personalized definitions for future use.

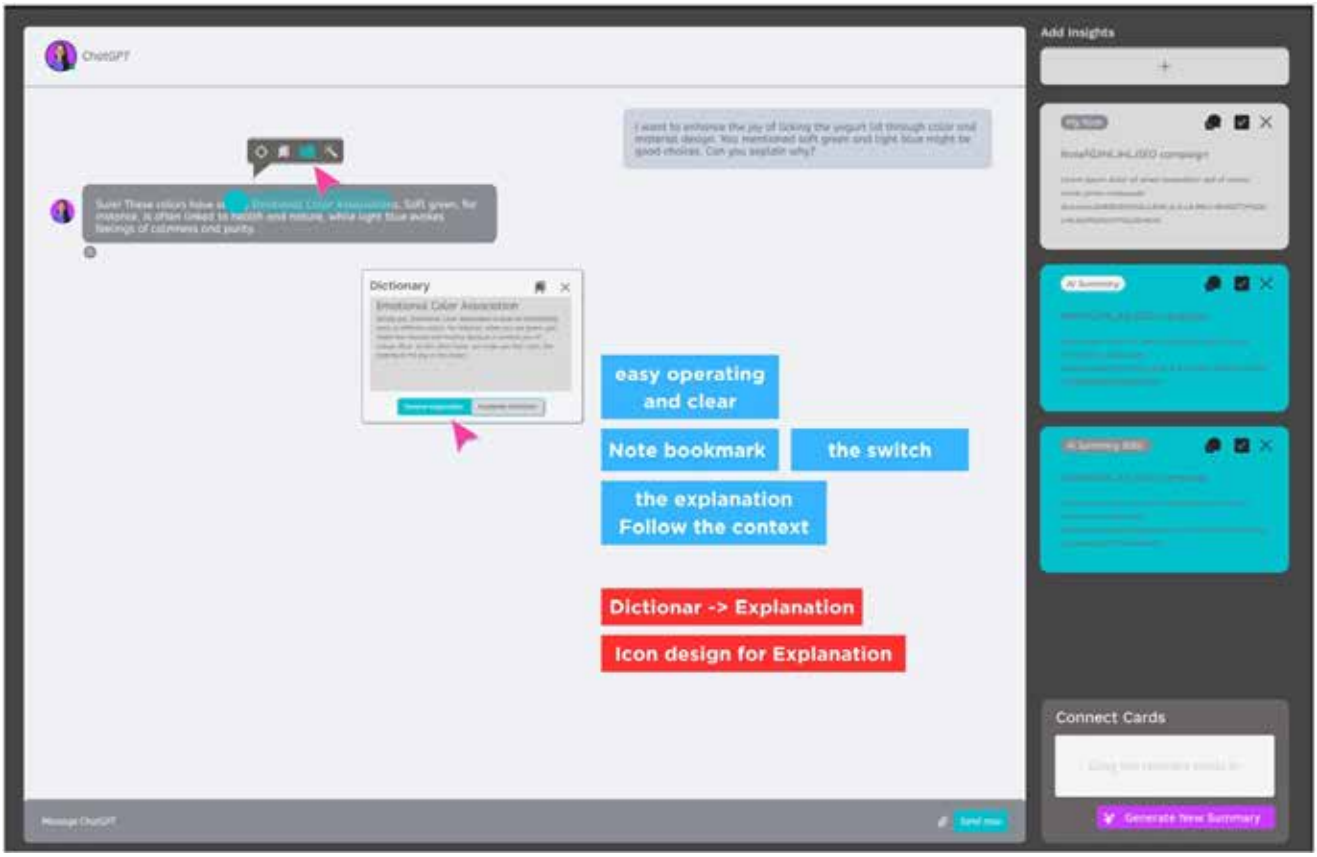
Discussion and Future Directions

The Dictionary feature demonstrates its effectiveness in reducing misunderstandings and improving communication clarity, aligning with the design goal of avoiding misleading communication. However, feedback on its limitations suggests room for improvement. Participants emphasized the importance of context when defining terms and expressed a need for functionality that integrates term explanations with the surrounding discussion.

To address these concerns, future iterations could incorporate an "explanation" feature that adapts definitions based on conversational context, allowing users to save customized definitions as bookmarks. Such enhancements would cater to designers' needs for term contextualization and facilitate more dynamic and meaningful discussions during collaborative tasks.

UI & UX Feedback:

Since explanations are based on the context of the conversation, users noted that the feature felt more like an "Explanation" tool. They suggested that an icon resembling a "magnifying glass" or a "question mark" would be more appropriate.



Function4: Card Note
Criteria: Align with users' real intentions, Be adjusted by the user

Align with users' real intentions

(V) Saving key conversation highlights

The ability to save key conversation highlights significantly aids users in recalling important points and organizing their thoughts. Participant A stated, *"After chatting with GPT for several days, I found certain arguments particularly valuable, so I recorded them in a notebook for later use."* This demonstrates that the Save Card feature effectively supports users in selecting and preserving critical information for future reference. By incorporating this functionality, the feature aligns with users' natural workflows, such as manually saving notes in external tools like Word, while streamlining the process within the interaction itself.

(V) Recalling context and improving information retrieval efficiency

The Track Back functionality addresses the common challenge of losing context in cumulative discussions. *Participant D explained, "When I revisit these cards, there are times I forget the context. Then I just click on this feature, and it takes me back to the original discussion."* This highlights how the feature reduces the cognitive load associated with recalling prior conversations and ensures seamless information retrieval. By providing quick access to past interactions, Track Back enhances users' ability to maintain conversational continuity, aligning with the goal of enabling users to clarify and pursue their real intentions efficiently.

(V) Enhancing information integration capabilities

Through card consolidation and summary generation, users can integrate diverse sources of information to form a more comprehensive understanding of topics. This functionality promotes synthesis and deeper analysis, facilitating effective problem-solving. Users highlighted how combining different cards helped them derive actionable insights, further demonstrating the feature's alignment with supporting users in achieving clarity and intentionality in their interactions.

(O) Supporting continuity of thought (Continuing thinking)

Users generally agreed that adding new cards supports fluid discussion and minimizes interruptions in thought processes. One participant noted, *"The card function allows me to continuously record information, maintaining the flow of my ideas."*

However, some users pointed out challenges in managing discussions with high information complexity. Participant D observed, *"Content can start pulling from earlier inputs or outputs, creating a mess that's hard to control. I often open new windows manually to separate discussions."* This feedback underscores the feature's partial success in facilitating continuous thinking, while also revealing its limitations in scenarios involving cumulative and intricate discussions.

Be adjusted by the user

(V) Compensating for ChatGPT's lack of critical thinking through interaction

The card-based interaction allows users to compensate for ChatGPT's limited critical thinking abilities. Participant C explained, *"I don't expect it to generate something with logical rigor. I write the arguments myself and let it supplement evidence and linking phrases."* This iterative process of creating and refining cards demonstrates the feature's capability to enable users to structure their thoughts and develop logical arguments more effectively.

(V) Classification and tagging tools

The feature also supports users in organizing saved cards through classification and tagging. Participant C shared, *"I create summaries by grouping cards based on my three key points." Another participant, B, noted, "*When I revisit old sessions, the cards are right there, making it easy to review past discussions." This highlights how the classification tools streamline information retrieval and reuse, enabling users to maintain clarity and focus across multiple conversations.

Summary and Future Improvements

The Save Card and Track Back features demonstrate strong alignment with the design goal of helping users clarify and achieve their real intentions. These functionalities empower users to save, revisit, and integrate information seamlessly, supporting more structured and intentional interactions. Additionally, the classification tools enhance organizational efficiency, catering to users' natural workflows.

However, some limitations remain in managing cumulative discussions and complex information. Users often resort to manual interventions, such as opening new windows, to maintain clarity and avoid content overlap. Future improvements could address these challenges by:

1. Enhancing contextual support: Allowing for dynamic topic separation and contextual alignment to reduce user effort in managing discussion complexity.
2. Streamlining tagging tools: Simplifying card classification and tagging processes to improve usability.
3. Integrating critical thinking aids: Providing enhanced support for logical structuring and argument development, reducing reliance on manual adjustments.

By addressing these areas, the system can further empower users in navigating complex tasks and fostering productive collaboration.

UI & UX Feedback:

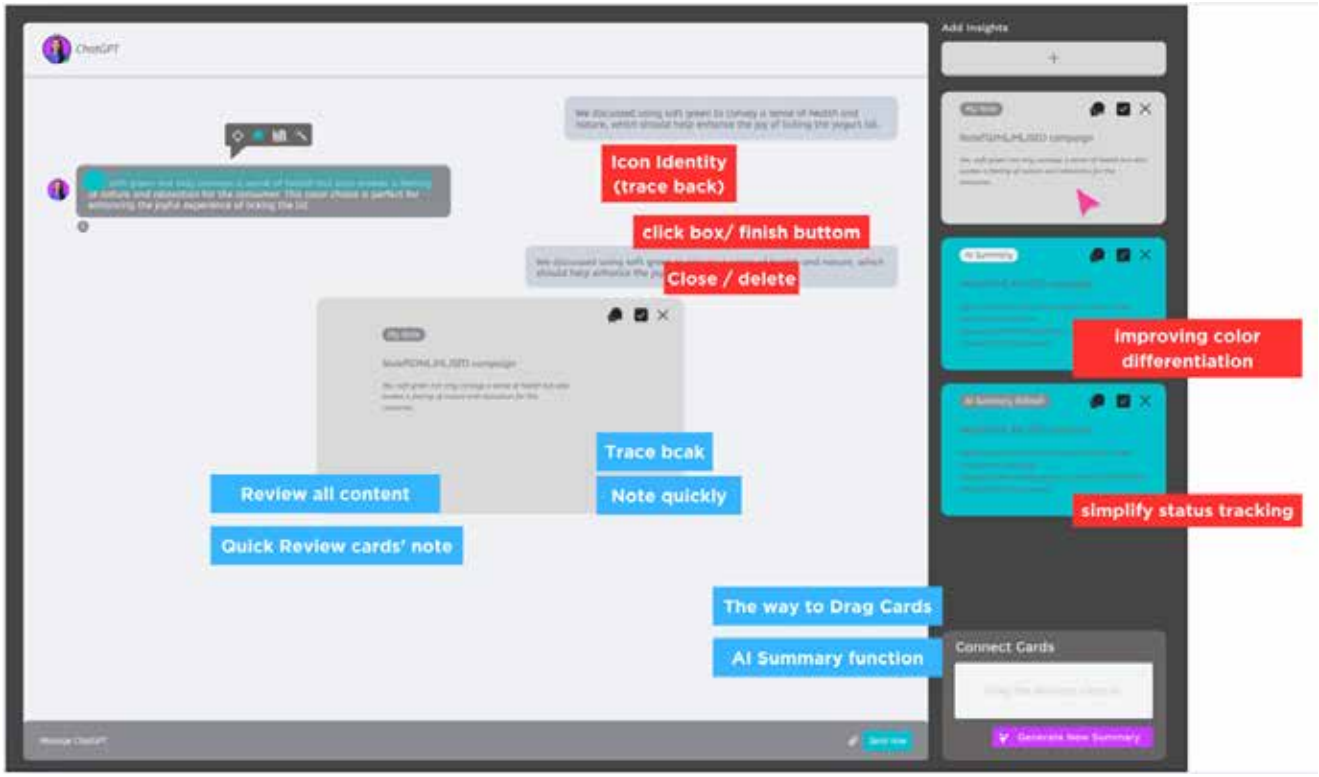
Visually, the "Bookmark & Note" feature is designed with three stages of card types:

1. My Note, 2. AI Summary, 3. AI Summary (Edited)

Users A, B, and D found the AI Summary to be a very convenient feature for consolidating numerous related cards. Since each AI Summary may not always be satisfactory, the option to edit the AI Summary as needed to add personal thoughts or refine the content was appreciated as a great design. However, User A mentioned that the colors used to display the three card types were not very memorable and suggested a simpler side-by-side layout for the three statuses (as shown in Figure 11) to make finding cards easier.

Users B and D had differing views on the "card merging pane." The former was "satisfied with the current size," while the latter suggested "enlarging it to the size of a card to facilitate the merging process or adding a checkbox feature with an extra merge button at the bottom." In short, the current window with the "Drag the relevant cards in" instruction effectively guides users to slide cards in for merging.

Lastly, participants noted that the icon buttons were somewhat misleading. For example, the X button at the top right of an expanded card could be interpreted as both a delete and a close window action. Additionally, the design of the backtracking button needed to be clearer. Currently, it is a dialogue bubble icon, but users felt it should resemble a "location" or "camera lock" icon.



6.3 Summary of Findings

The evaluation of the features—Off-Topic, Prompt Guide, Dictionary, and CardNote—against the three design criteria (Align with Users' Real Intentions, Avoid Misleading Communication, and Be Adjusted by the user) highlights their strengths in supporting user needs and identifies areas for improvement.

Summary (Table)
Align with Users' Real Intentions
(V): All features strongly align with helping users clarify goals and organize information.
(O): High-complexity scenarios reveal gaps in scalability.
Avoid Misleading Communication
(V): Dictionary and CardNote excel in preventing misunderstandings.
(X): Prompt Guide lacks specificity in advanced contexts.
Be Adjusted by the User
(V): Features like Off-Topic and CardNote offer intuitive customization.
(O): Automation is needed for large information volumes.

Align with Users' Real Intentions

All four features demonstrated strong alignment to help users clarify their goals and organize information. Off-Topic empowers users to focus on key conversation topics by filtering out distractions, enabling smoother and more productive interactions. Similarly, CardNote supports users in saving, organizing, and synthesizing critical information, facilitating clarity and intentionality during discussions.

However, both features exposed gaps in scalability for both features in high-complexity scenarios. For instance, while Off-Topic excels in managing small-to-moderate amounts of content, its reliance on manual adjustments, such as dragging and dropping, becomes inefficient when handling large volumes of information. Similarly, CardNote struggles with cumulative discussions where the overlap of prior inputs or outputs creates confusion, requiring users to reorganize or separate content for clarity manually.

Avoid Misleading Communication

Features like Dictionary and CardNote prevent misunderstandings, aligning strongly to reduce ambiguity. Dictionary provides users with straightforward definitions of unfamiliar terms, enabling smoother communication. Users highlighted its ability to clarify technical terms and reduce cognitive load, particularly when faced with specialized language. Meanwhile, CardNote's *Track Back* functionality ensures that users can easily recall the prior context in cumulative discussions, reducing the likelihood of losing track of the conversation and maintaining clarity over time.

However, Prompt Guide lacked specificity, particularly in advanced contexts where users require more tailored suggestions. Generalized outputs sometimes cause the AI to deviate from the user's original intention, leading to inefficiencies. Experienced users expressed frustration with the lack of precision, suggesting a mismatch between the feature's design and the needs of those engaging in more complex or iterative tasks.

Be Adjusted by the user

Features like Off-Topic and CardNote demonstrated strong alignment with this criterion by offering intuitive tools for user customization. For example, Off-Topic's drag-and-drop mechanism allows users to adjust conversation direction seamlessly, while CardNote's classification and tagging tools enable users to organize information in ways that suit their personal workflows. These features empower users to take control of their interactions, promoting flexibility and usability.

However, automation is needed for large information volumes. In both Off-Topic and CardNote, users highlighted the challenge of manually handling and organizing complex discussions or extensive content. For instance, in Off-Topic, participants suggested integrating AI-driven prioritization tools to detect and filter key points automatically. Similarly, in CardNote, participants noted that cumulative content management often required opening new windows or separating discussions manually, creating unnecessary friction.

In summary, all features align strongly with their intended goals but require refinements to address scalability and specificity in advanced contexts. Features like Dictionary and CardNote excel in clarity and context retention, while Off-Topic and Prompt Guide provide valuable user control and ideation tools. However, introducing automation and context-aware functionality will be critical to overcoming the identified limitations and enhancing the overall user experience.

7

Recommendation & Conclusion

7.1 Future Recommendation

To address the identified limitations and enhance the features' effectiveness, the following recommendations are proposed:

1. Introduce Dynamic Contextual Integration

Features like Prompt Guide and Dictionary should incorporate context-aware functionality to adapt dynamically to ongoing conversations. For instance, Prompt Guide could refine its suggestions based on the user's intent and prior interactions, ensuring greater specificity and relevance. Similarly, a dictionary could provide definitions tailored to the context of the discussion, reducing the need for generic explanations.

2. Develop Scalable Automation Tools

High-information scenarios expose the inefficiency of manual adjustments in Off-Topic and CardNote. AI-driven tools should be integrated to automate content prioritization, filtering, and organization. For example, Off-Topic could include an automated key point detection feature, while CardNote could benefit from dynamic clustering to group related cards based on user-defined criteria.

3. Enhance User-Driven Customization

Users consistently highlighted the need for more outstanding customization options. Features like Dictionary allow users to save and edit personalized definitions for recurring use. At the same time, Prompt Guide offers adjustable levels of suggestion specificity to cater to novice and advanced users.

4. Implement Interactive Visual Feedback

All features could benefit from intuitive visual cues like color coding, animations, and dynamic indicators. For instance, CardNote could visually highlight frequently revisited cards, and Off-Topic could use animations to show how content adjustments impact conversation flow. These enhancements would reinforce usability and provide immediate feedback to the user.

7.2 Limitation

The evaluation of the features identified several limitations that constrain their current effectiveness and generalizability:

1. Scalability in High-Complexity Scenarios

While features like Off-Topic and CardNote excel in managing small-to-moderate amounts of information, they struggle in high-information contexts. Users found manual adjustments, such as dragging and tagging, could be more efficient when dealing with large volumes of data or cumulative discussions. This limitation reduces the features' usability in more complex real-world applications.

2. Limited Contextual Adaptability

Features like Prompt Guide and Dictionary lack the ability to adapt dynamically to conversation context. Generic outputs and definitions often fail to align with user-specific needs, particularly in advanced or iterative tasks. This limitation highlights a critical gap between the current feature design and the nuanced requirements of experienced users.

3. Sample Size and Use Case Specificity

The study was conducted with a relatively small and specific participant group, limiting the generalizability of findings. Feedback was gathered under controlled use-case scenarios, such as brainstorming or content organization, which may not fully capture the diverse range of potential user interactions in real-world contexts.

4. Dependency on User Expertise

Several features, such as Prompt Guide and Dictionary, rely heavily on users' ability to effectively frame prompts or interpret outputs. While these tools provide valuable support for beginners, they may feel redundant or cumbersome for experienced users who prefer direct control over their interactions.

5. Focus on Interaction Rather Than Outcomes

The evaluation prioritized user interaction with the features, such as their ease of use and alignment with design criteria, rather than long-term outcomes like task efficiency or productivity gains. As a result, the impact of these features on overall user performance still needs to be explored.

By addressing these limitations and incorporating the proposed recommendations, the features can evolve to meet a broader user needs, ensuring scalability, adaptability, and effectiveness in diverse and dynamic contexts.

References

1. Umbrello, S., Capasso, M., Balistreri, M., Pirni, A., & Merenda, F. (2021). Value sensitive design to achieve the UN SDGs with AI: A case of elderly care robots. Minds and Machines, 31(3), 395-419.
2. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds & Machines, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5
3. Umbrello, S., & Yampolskiy, R. V. (2022). Designing AI for explainability and verifiability: a value sensitive design approach to avoid artificial stupidity in autonomous vehicles. International Journal of Social Robotics, 14(2), 313-322.
4. Guilherme, A. (2019). AI and education: the importance of teacher and student relations. AI & society, 34, 47-54.
5. Schiff, D., Biddle, J., Borenstein, J., & Laas, K. (2020b). What's next for AI ethics, policy, and governance? A global overview. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 153–58. AIES '20. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3375627.3375804.
6. Schiff, D., Borenstein, J., Laas, K., & Biddle, J. (2021). AI ethics in the public, private, and NGO sectors: A review of a global document collection. IEEE Transactions on Technology and Society, 2, 31–42. https://doi.org/10.1109/TTS.2021.3052127
7. Holmes, W., & Porayska-Pomsta, K. (2018). The ethics of artificial intelligence in education. University Business.
8. Ali, S.; Payne, B.H.; Williams, R.; Park, H.W.; Breazeal, C. Constructionism, ethics, and creativity: Developing primary and middle school artificial intelligence education. In Proceedings of the International Workshop on Education in Artificial Intelligence K-12 (Eduai'19), Macao, China, 10–16 August 2019; pp. 1–4.
9. Su, J.; Zhong, Y.; Ng, D.T.K. A meta-review of literature on educational approaches for teaching AI at the K-12 levels in the Asia-Pacific region. Comput. Educ. Artif. Intell. 2022, 3, 100065. [CrossRef]
10.Touretzky, D.; Gardner-McCune, C.; Martin, F.; Seehorn, D. Envisioning AI for K-12: What should every child know about AI? In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; pp. 9795–9799.
11. Ryan, R.M., and E.L. Deci. 2000. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. The American Psychologist 55: 68–78. https://doi.org/10.1037/0003-066X.55.1.68.
———. 2017. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness. New York: Guilford Press.
12. Ryan, R.M. 1982. Control and Information in the Intrapersonal Sphere: An Extension of Cognitive Evaluation Theory. Journal of Personality and Social Psychology 43 (3): 450.
13. Christman, J., ed. 1989. The Inner Citadel: Essays on Individual Autonomy. New York: Oxford University Press.
———. Autonomy in Moral and Political Philosophy. The Stanford Encyclopedia of Philosophy (Spring 2018 Edition), ed. Edward N. Zalta, Available online https://plato.stanford.edu/archives/spr2018/entries/autonomy-moral/.
14. Przybylski, A.K., N. Weinstein, R.M. Ryan, and C.S. Rigby. 2009. Having to Versus Wanting to Play: Background and Consequences of Harmonious Versus Obsessive Engagement in Video Games. Cyber Psychology & Behavior 12 (5): 485–492. https://doi.org/10.1089/cpb.2009.0083.
15. Przybylski, A.K., K. Murayama, C.R. Dehaan, and V. Gladwell. 2013. Motivational, Emotional, and Behavioral Correlates of Fear of Missing Out. Computers in Human Behavior. https://doi.org/10.1016/j.chb.2013.02.014.

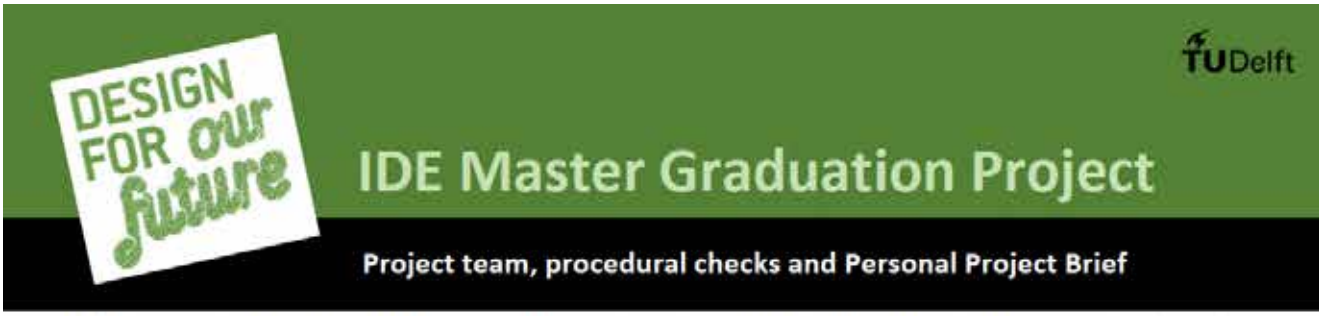
16. Hill, T. 2013. Kantian Autonomy and Contemporary Ideas of Autonomy. In Kant on Moral Autonomy, ed. Oliver Sensen, 15–31. Cambridge: Cambridge University Press.
17. Siegel, M. The sense-think-act paradigm revisited. In Proceedings of the 1st International Workshop on Robotic Sensing, Örebro, Sweden, 5–6 June 2003.
18. Hayles, N.K. Computing the human. Theory Cult. Soc. 2005, 22, 131–151. [CrossRef]
19. Beer, J.M.; Fisk, A.D.; Rogers, W.A. Toward a framework for levels of robot autonomy in human-robot interaction. J. Hum.-Robot. Interact. 2014, 3, 74. [CrossRef]
20. DIXIT Cards. (2014). In Wikipedia. Retrieved from [URL of the Wikipedia page]
21. Kwiatkowska, J., Szóstek, A., & Lamas, D. (2014, October). (Un) structured sources of inspiration: comparing the effects of game-like cards and design cards on creativity in co-design process. In Proceedings of the 13th Participatory Design Conference: Research Papers-Volume 1 (pp. 31-39).
22. Sousa, M. (2020, May). Fast Brainstorm techniques with modern board games adaptations for daily uses in business and project managing. In Proceedings of the International Conference of Applied Business and Manage-ment (ICABM2020) (pp. 508-524).
23. Van de Poel, I. (2013). Translating values into design requirements. In Philosophy of engineering and technology (pp. 253–266, Vol. 15). Springer Science+Business.
24. Schiff, D. (2022). Education for AI, not AI for education: The role of education and ethics in national AI policy strategies. International Journal of Artificial Intelligence in Education, 32(3), 527-563.
25. Ministry of the Economy and Innovation. (2019). Lithuanian Artificial Intelligence Strategy: A Vision of the Future. Vilnius, Lithuania: Ministry of the Economy and Innovation. http://kurklt.lt/wp-conte nt/uploads/2018/09/StrategyIndesignpdf.pdf.
26. Smart Nation and Digital Government Office, Singapore. (2019). Singapore: National Artificial Intelligence Strategy: Advancing Our Smart Nation Journey. Singapore: Smart Nation and Digital Government Office. https://www.smartnation.sg/docs/default-source/default-document-library/national-ai-strategy. pdf.
27. NITI Aayog, India. (2018). India: National Strategy for Artificial Intelligence #AIFORALL. Delhi, India: NITI Aayog. http://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI- Discussion-Paper.pdf.
28. Office of the Prime Minister, Malta. (2019). Malta: Towards An AI Strategy. Valletta, Malta: Parliamen- tary Secretariat for Financial Services, Digital Economy and Innovation, Office of the Prime Min- ister, Malta. http://malta.ai/wp-content/uploads/2019/04/Draft_Policy_document_-_online_versi on.pdf.
29. Floridi, L., Cowls, J., King, T. C., & Taddeo, M. (2021). How to design AI for social good: Seven essential factors. Ethics, Governance, and Policies in Artificial Intelligence, 125-151.
30. Hussain, A. (2019). PERSONAL INTEGRATIVE NEEDS AND USE OF SOCIAL MEDIA AMONG INFORMATION PROFESSIONALS.
31. Friedman, B., Kahn, P. H., Jr., Borning, A., & Huldtgren, A. (2013). Value sensitive design and information systems. In N. Doorn, D. Schuurbiers, I. van de Poel, & M. E. Gorman (Eds.), Early engagement and new technologies: Opening up the laboratory (pp. 55–95). Springer Netherlands. https://doi.org/10.1007/978-94-007-7844-3_4
32. van de Poel, I. (2013). Translating values into design requirements. In D. P. Michelfelder, N. McCarthy, & D. E. Goldberg (Eds.), Philosophy and engineering: Reflections on practice, principles and process (pp. 253–266). Springer Netherlands. https://doi.org/10.1007/978-94-007-7762-0_20
33. Umbrello, S., & van de Poel, I. (2021). Mapping value sensitive design onto AI for social good principles. AI and Ethics. https://doi.org/10.1007/s43681-021-00038-3

Appendix A: Project Brief

Appendix B: Workshop analysis- All broken down needs and pains

Appendix C: Code book

Appendix A: Project Brief



In this document the agreements made between student and supervisory team about the student's IDE Master Graduation Project are set out. This document may also include involvement of an external client, however does not cover any legal matters student and client (might) agree upon. Next to that, this document facilitates the required procedural checks:

- Student defines the team, what the student is going to do/deliver and how that will come about
- Chair of the supervisory team signs, to formally approve the project's setup / Project brief
- SSC E&SA (Shared Service Centre, Education & Student Affairs) report on the student's registration and study progress
- IDE's Board of Examiners confirms the proposed supervisory team on their eligibility, and whether the student is allowed to start the Graduation Project

STUDENT DATA & MASTER PROGRAMME
Complete all fields and indicate which master(s) you are in

Family name	Yang	7175	
Initials	H.R		
Given name	Hui-Ru		
Student number	5122880		

IDE master(s) IPD ☐ Dfl ✓ SPD ☐

2nd non-IDE master

Individual programme *(date of approval)*

Medisign ☐

HPM ☐

SUPERVISORY TEAM
Fill in he required information of supervisory team members. If applicable, company mentor is added as 2nd mentor

Chair	Sara Colombo	dept./section	SDE/ HCAI
mentor	Eric Heng GU	dept./section	DOS/ MOD
2nd mentor			
client:			
city:		country:	
optional comments			

! Ensure a heterogeneous team. In case you wish to include team members from the same section, explain why.

! Chair should request the IDE Board of Examiners for approval when a non-IDE mentor is proposed. Include CV and motivation letter.

! 2nd mentor only applies when a client is involved.

APPROVAL OF CHAIR on PROJECT PROPOSAL / PROJECT BRIEF -> to be filled in **by the Chair** of the supervisory team

Sign for approval (Chair)

Sara Colombo — Digitally signed by Sara Colombo Date: 2024.05.28 12:02:09 +02'00'

Name	Sara Colombo	Date	28 May 2024	Signature	

CHECK ON STUDY PROGRESS

To be filled in **by SSC E&SA** (Shared Service Centre, Education & Student Affairs), after approval of the project brief by the chair.
The study progress will be checked for a 2nd time just before the green light meeting.

Master electives no. of EC accumulated in total		EC
Of which, taking conditional requirements into account, can be part of the exam programme		EC

★	YES	all 1st year master courses passed
	NO	missing 1st year courses

Comments:

Sign for approval (SSC E&SA)

Rik Ledoux Digitally signed by Rik Ledoux Date: 2024.05.24 11:03:58 +02'00'

Name Rik Ledoux Date 24 May 2024 Signature

APPROVAL OF BOARD OF EXAMINERS IDE on SUPERVISORY TEAM -> to be checked and filled in by IDE's Board of Examiners

Does the composition of the Supervisory Team comply with regulations?

YES	★	Supervisory Team approved
NO		Supervisory Team not approved

Comments:

Based on study progress, students is ...

★	**ALLOWED** to start the graduation project
	NOT allowed to start the graduation project

Comments:

Sign for approval (BoEx)

Monique von Morgen Digitally signed by Monique von Morgen Date: 2024.05.28 11:27:15 +02'00'

Name Monique von Morgen Date 28 May 2024 Signature

Personal Project Brief – IDE Master Graduation Project

Name student Hui-Ru, Yang **Student number** 5,122,880

PROJECT TITLE, INTRODUCTION, PROBLEM DEFINITION and ASSIGNMENT

Complete all fields, keep information clear, specific and concise

Project title Shaping Ethical Interactions: A Value Sensitive Design Lens on AI Agent and User Autonomy

Please state the title of your graduation project (above). Keep the title compact and simple. Do not use abbreviations. The remainder of this document allows you to define and clarify your graduation project.

Introduction

Describe the context of your project here; What is the domain in which your project takes place? Who are the main stakeholders and what interests are at stake? Describe the opportunities (and limitations) in this domain to better serve the stakeholder interests. (max 250 words)

AI agents are being used to interact with human beings in various educational scenarios. Chatbots and related algorithms can simulate human interactions and generate human-like text based on natural language input (Adiguzel, 2023). For instance, students use ChatGPT for various academic tasks, including coding guidance, writing support, and fostering and conceptualizing ideas. Furthermore, ChatGPT affords one-on-one communicative capabilities as a virtual tutor for instructors, offering valuable support in tasks such as generating course materials and providing suggestions that effectively address individual students' specific queries.

However, a sense of convenience in its users also contributes to users' dependency. This overreliance raises concerns about the potential erosion of human autonomy. Floridi and Cowls (2019) discussed the concept of autonomy, where humans transfer decision-making authority to artificial intelligence. They express concerns that "the growth in artificial autonomy may undermine the flourishing of human autonomy" (Floridi & Cowls, 2019, p. 7). Humans think using AI in their decisions could benefit them and save time. However, it overcomes human biological processors by lowering cognition capabilities (Jarrahi, 2018). Research has pointed out that the potential consequences of using AI models in academic writing may manifest in a decline in critical thinking abilities and a decrease in innovative thinking among students and teachers (Ahmad et al., 2023). This paper will research how to empower students' autonomy when using ChatGPT for their academic tasks by applying the VSD framework.

→ *space available for images / figures on next page*

introduction (continued): space for images

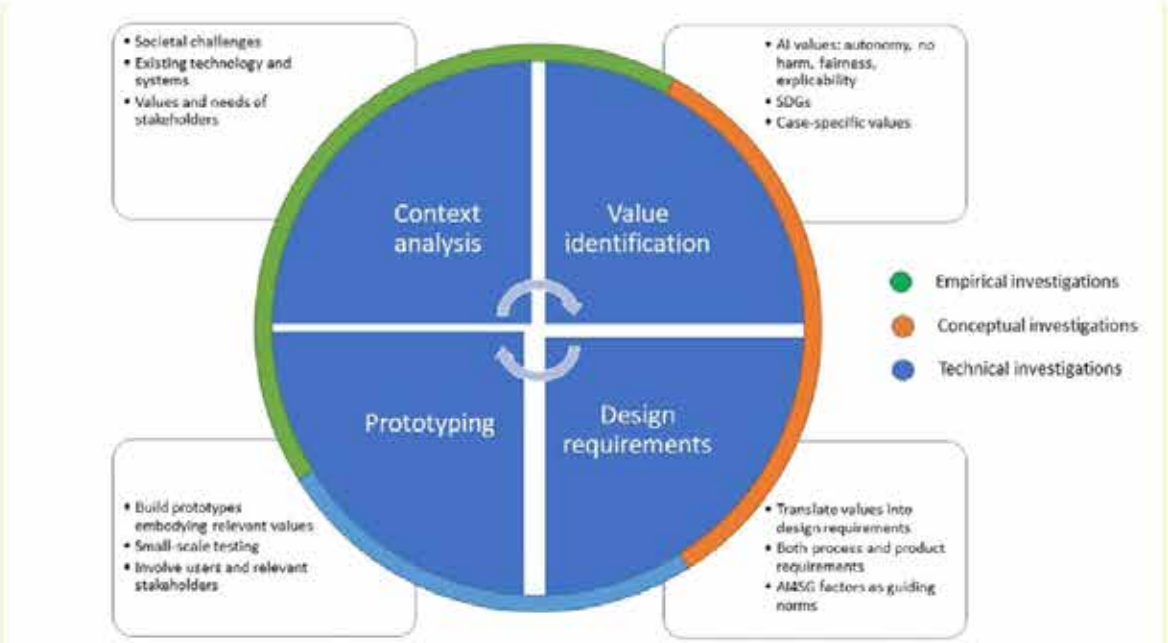


Fig. 6 VSD design process for AI technologies

image / figure 1 VSD design process for AI technologies

image / figure 2 Gannt Chart

DESIGN FOR our future

TUDelft

Personal Project Brief – IDE Master Graduation Project

Problem Definition

What problem do you want to solve in the context described in the introduction, and within the available time frame of 100 working days? (= Master Graduation Project of 30 EC). What opportunities do you see to create added value for the described stakeholders? Substantiate your choice.
(max 200 words)

This study aims to research how AI models impact the autonomy of college students aged 20 to 27 in their learning experiences, particularly when the interactions and mindset/behavior shift.

Additionally, the study will explore how the value-sensitive design approach could empower AI models' ethical values, such as ChatGPT, and influence students' ability to empower their' autonomy and foster a supportive learning environment.

Below are the specific research objectives.
RQ1. Determine stakeholders' change in mindset and behavior when interacting with ChatGPT on academic tasks.
RQ2. Determine the stakeholders' autonomy values in education when using ChatGPT for their academic tasks.
RQ3. Determine how the value-sensitive design approach could empower AI models' autonomy values.

These objectives seek to provide insights into how AI models can promote students' autonomy in the educational context. Ultimately, the study aims to develop prototypes for new features to embed in future designs for AI agents in higher education.

Assignment

This is the most important part of the project brief because it will give a clear direction of what you are heading for. Formulate an assignment to yourself regarding what you expect to deliver as result at the end of your project. (1 sentence) As you graduate as an industrial design engineer, your assignment will start with a verb (Design/Investigate/Validate/Create), and you may use the green text format:

I want to design prototypes as a case study to execute and analyze the application of the Value-Sensitive Design framework f or ethical AI design to enhance ethical user-AI interactions in AI agent contexts.

Then explain your project approach to carrying out your graduation project and what research and design methods you plan to use to generate your design solution (max 150 words)

This research employs a Value-Sensitive Design (VSD) approach to complement a design thinking approach, facilitating more comprehensive and profound design solutions. It is divided into four main phases: context analysis, value identification, desi gn requirements, and prototyping.
This project only focuses on the AI agent field as a case study. It will include a literature review, desk research, surveys, and i nterviews with AI-design experts and users to understand the latent concerns and needs. Stakeholder values and user journe y maps will be created to showcase the boundaries and opportunities in the design process by applying the VSD framework. Next, the prototypes will be designed and evaluated during this graduation project. Finally, a few analyses will be run to vali date the VSD framework for the insights and effects.

Project planning and key moments

*To make visible how you plan to spend your time, you must make a planning for the full project. You are advised to use a Gantt chart format to show the different phases of your project, deliverables you have in mind, meetings and in-between deadlines. Keep in mind that all activities should fit within the given run time of 100 working days. Your planning should include a **kick-off meeting**, **mid-term evaluation meeting**, **green light meeting** and **graduation ceremony**. Please indicate periods of part-time activities and/or periods of not spending time on your graduation project, if any (for instance because of holidays or parallel course activities).*

Make sure to attach the full plan to this project brief.
The four key moment dates must be filled in below

Kick off meeting	19 Mar 2024
Mid-term evaluation	11 Jun 2024
Green light meeting	12 Jul 2024
Graduation ceremony	26 Aug 2024

In exceptional cases (part of) the Graduation Project may need to be scheduled part-time. Indicate here if such applies to your project

Part of project scheduled part-time	
For how many project weeks	
Number of project days per week	

Comments:

Motivation and personal ambitions

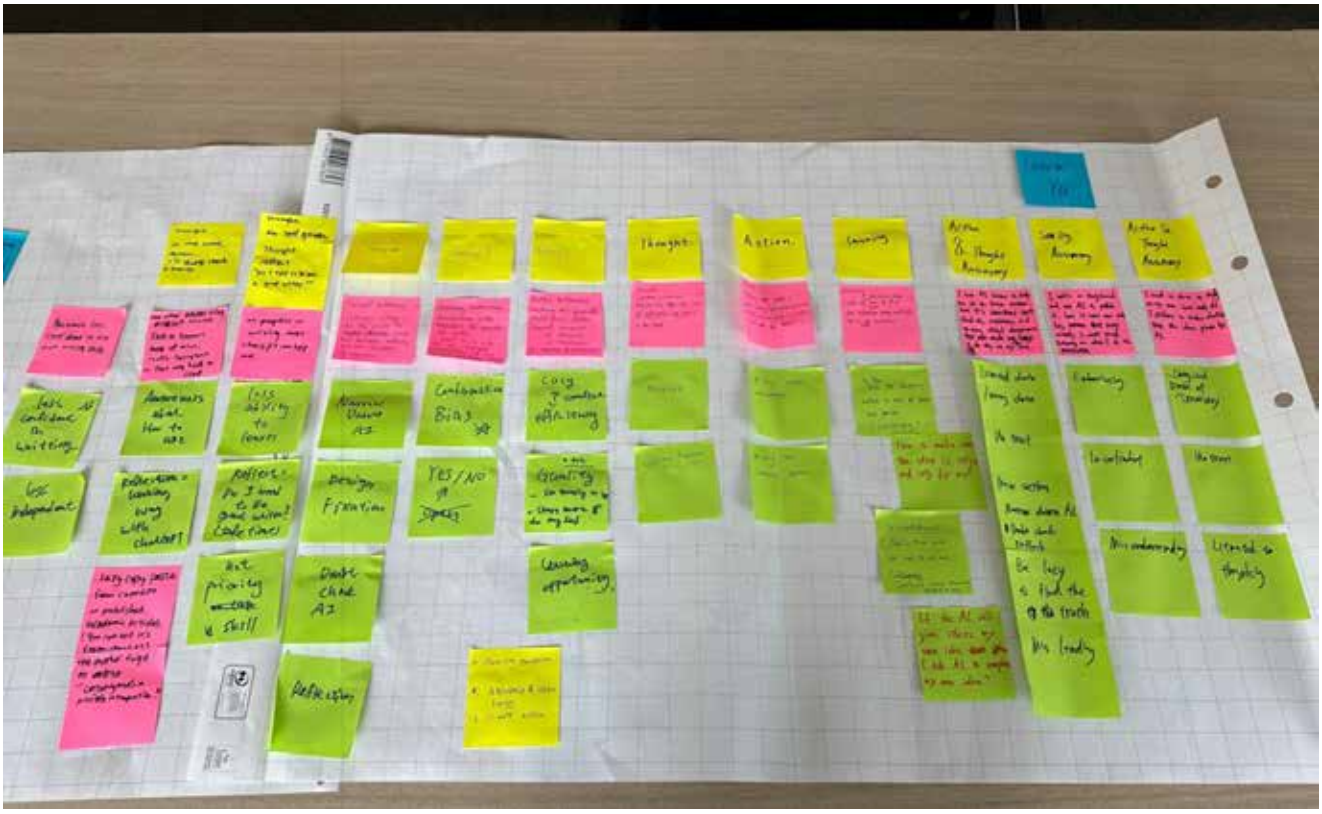
Explain why you wish to start this project, what competencies you want to prove or develop (e.g. competencies acquired in your MSc programme, electives, extra-curricular activities or other).

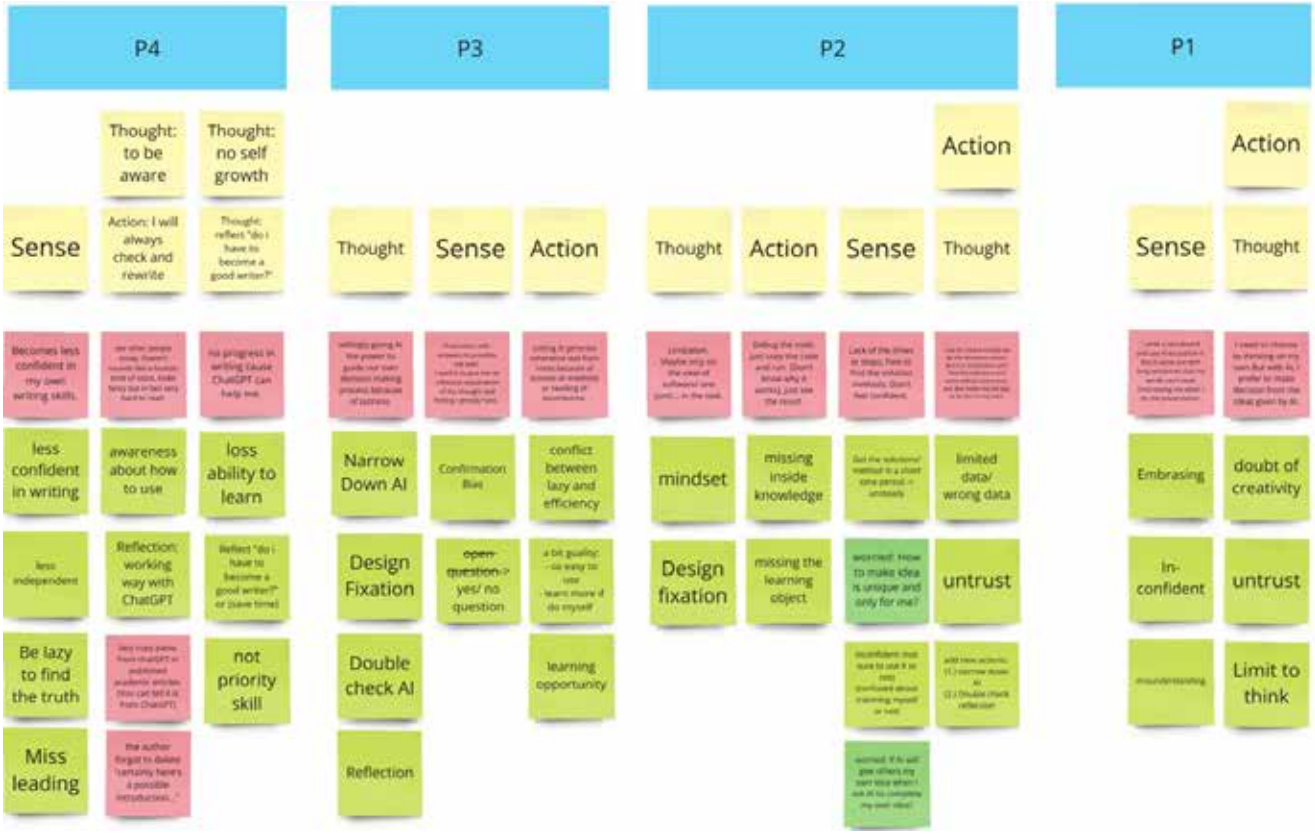
Optionally, describe whether you have some personal learning ambitions which you explicitly want to address in this project, on top of the learning objectives of the Graduation Project itself. You might think of e.g. acquiring in depth knowledge on a specific subject, broadening your competencies or experimenting with a specific tool or methodology. Personal learning ambitions are limited to a maximum number of five.
(200 words max)

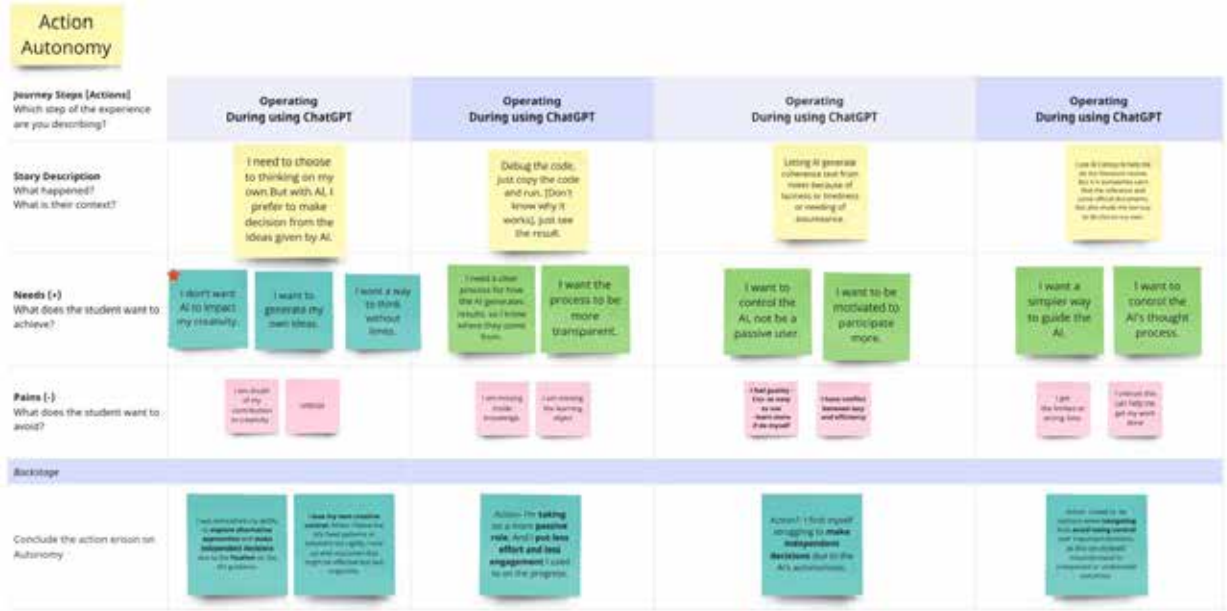
I gained industry and academic knowledge of AI during my previous work and study. I became passionate about technology's potential development but always needed help with the gap between thinking and doing. Thus, in this graduation project, I expect to improve my knowledge and communication skills to design AI products for human benefit.
There will be opportunities to interview and workshop with stakeholders during the process. I am ambitious about finding th e gap between theory and practice.

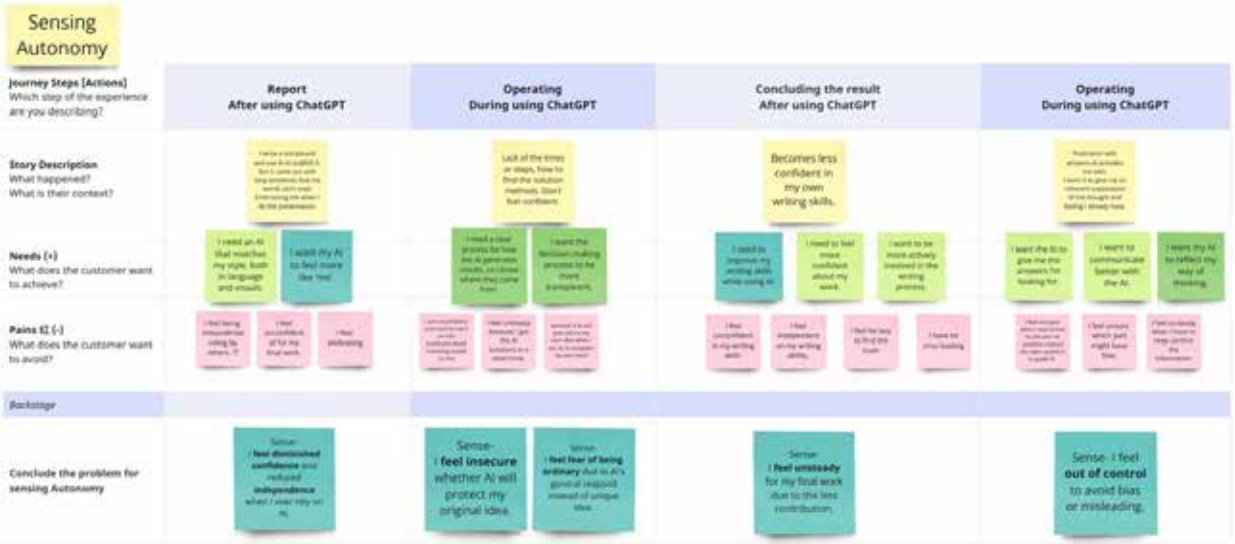
My Intended Outcome would be:
1. Prototypes: Develop prototypes as a case study for redesigned AI agent interfaces and interactions, focusing on applying t he VSD framework for ethical AI design (Umbrello & van de Poel, 2021). These examples will showcase practical ways to shap e ethical interactions between users and AI.
2. Impact Analysis: Based on VSD principles, perform an in-depth impact analysis on the redesigned AI agent. This analysis wi ll explore changes in user perceptions regarding autonomy and dependency, aiming to provide empirical insights into the eff ects of ethical redesign practices in AI agents.

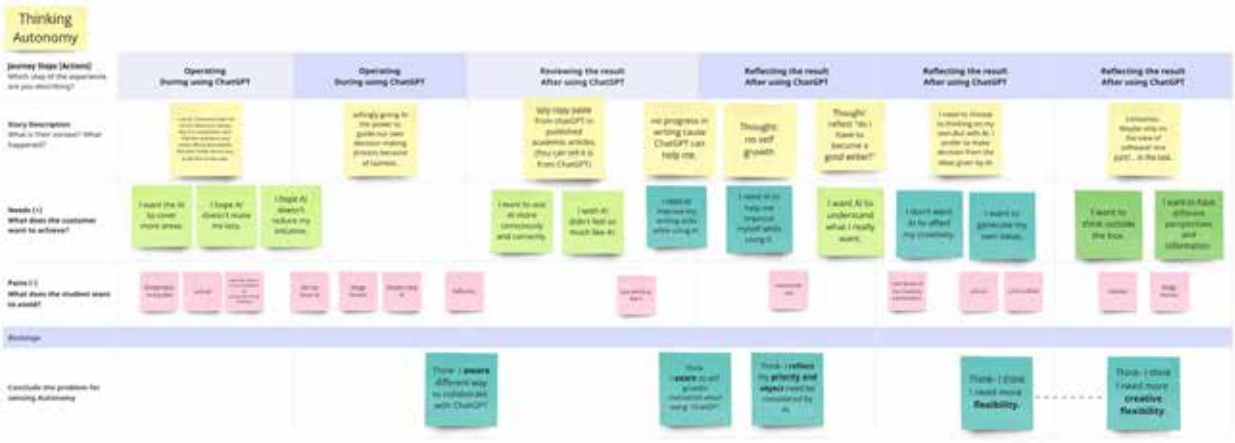
Appendix B: Workshop analysis- All broken down needs and pains






Action Autonomy
Journey Steps [Actions]
Which step of the experience are you describing?
Operating During using ChatGPT
Operating During using ChatGPT
Operating During using ChatGPT
Operating During using ChatGPT
Story Description
What happened?
What is their context?
I need to choose to thinking on my own.But with AI, I prefer to make decision from the ideas given by AI.
Debug the code, just copy the code and run. [Don't know why it works], just see the result.
Letting AI generate coherence text from notes because of laziness or needing of assistance
Needs (+)
What does the student want to achieve?
I don't want AI to impact my creativity.
I want to generate my own ideas.
I want a way to think without limits.
I want the process to be more transparent.
I want to control the AI, not be a passive user.
I want to be motivated to participate more.
I want a simpler way to guide the AI.
I want to control the AI's thought process.
Pains (-)
What does the student want to avoid?
Backstage
Conclude the action ersion on Autonomy


Sensing Autonomy
Journey Steps [Actions]
Which step of the experience are you describing?
Report After using ChatGPT
Operating During using ChatGPT
Concluding the result After using ChatGPT
Operating During using ChatGPT
Story Description
What happened?
What is their context?
Becomes less confident in my own writing skills.
Needs (+)
What does the customer want to achieve?
I need an AI that matches my style, both in language and emotio.
I want my AI to feel more like 'me'.
I need a clear process for how the AI generation results, so I know where they come from.
I want the decision-making process to be more transparent.
I want to be more actively involved in the writing process.
I want the AI to give me the answers I'm looking for.
I want to communicate better with the AI.
I want my AI to reflect my way of thinking.
Pains (-)
What does the customer want to avoid?
Backstage
Conclude the problem for sensing Autonomy
Sense- I feel diminished confidence and reduced independence when I over rely on AI.
Sense- I feel insecure whether AI will protect my original idea.
Sense- I feel fear of being ordinary due to AI's general respond instead of unique idea.
Sense- I feel unsteady for my final work due to the less contribution.
Sense- I feel out of control to avoid bias or misleading.


Thinking Autonomy
Journey Steps [Actions]
Which step of the experience are you describing?
Operating During using ChatGPT
Operating During using ChatGPT
Reviewing the result After using ChatGPT
Reflecting the result After using ChatGPT
Reflecting the result After using ChatGPT
Reflecting the result After using ChatGPT
Story Description
What is their context? What happened?
Needs (+)
What does the customer want to achieve?
I want the AI to cover more areas
I hope AI doesn't make me lazy.
I hope AI doesn't reduce my initiative.
I want to use AI more consciously and correctly.
I want AI didn't feel so much like AI.
I want to improve my writing skills when using AI.
I want to help me improve myself while using it.
I want AI to understand what I really want.
I don't want AI to affect my creativity.
I want to generate my own ideas.
I want to think outside the box.
I want to have different perspectives and information
Pains (-)
What does the student want to avoid?
Backstage
Conclude the problem for sensing Autonomy
Think- I aware different way to collaborate with ChatGPT
Think- I aware AI and growth (motivation when using ChatGPT)
Think- I reflect my priority and object need be considered by AI.
Think- I think I gain flexibility.
Think- I think I need more creative flexibility.

Appendix C: Code book

Theme	Code	Groundedness
Align with users' real intentions,	(V) Helping users focus on key conversation topics	6
	(V) Enhanced engagement in conversations	8
	(V) Smoother interactions (V) Improving clarity in expressing user intentions	5
	(V) Supporting brainstorming to inspire ideas	6
	(O) Reducing trial-and-error inputs	4
	(X) Lack of specificity in prompt suggestions	3
	(V) Saving key conversation highlights	14
	(V) Recalling context and improving information retrieval efficiency	13
	(V) Enhancing information integration capabilities	4
	(O) Supporting continuity of thought (Continuing thinking)	8
	Inspiring	6
	Interactive Features	4
	Misunderstanding	2
	Organizing thoughts or ideas	3
	Reminders/ tutorial for users	1
	User choice	1
Be adjusted by the user	(V) Flexibly adjusting the conversation direction	8
	(V) Reducing distractions from irrelevant content	9
	(X) The need for more efficient handling of large volumes of content	10
	(V) Compensating for ChatGPT's lack of critical thinking through interaction	13
	(V) Classification and tagging tools	16
	Reflection	2
	Awareness	3
	Lack of Critical thinking and Creativity	2
	Collaboration with ChatGPT	5
	Control	2
	Involvement	2
Avoid misleading communication	(V) Reducing misunderstandings caused by technical terms	9
	(X) Limited contextual adaptability	12
	Functionality	16
	Inspiration	13
	Communication	8
	Purpose	2
	No Trust	1
	Adaptability	2
	Autonomy	6
	Convenience	1
	Novelty	4