



CHALMERS
UNIVERSITY OF TECHNOLOGY



UNIVERSITY OF GOTHENBURG

Incorporating AI into the UX Design Process

A Study of Concept Exploration and Ideation

Master's thesis in Computer science and engineering

KIMIA MOHAMMADI MOHAMMADI, ELIN RUUD

MASTER'S THESIS 2026

Incorporating AI into the UX Design Process

A Study of Concept Exploration and Ideation

KIMIA MOHAMMADI MOHAMMADI, ELIN RUUD



UNIVERSITY OF
GOTHENBURG



CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2026

Incorporating AI into the UX Design Process
A Study of Concept Exploration and Ideation
KIMIA MOHAMMADI MOHAMMADI, ELIN RUUD

© KIMIA MOHAMMADI MOHAMMADI, ELIN RUUD, 2026.

Supervisor: Martina De Cet, Department of Computer Science and Engineering
Co-supervisor: Mohammad Obaid, Department of Computer Science and Engineering
Examiner: Sara Ljungblad, Department of Computer Science and Engineering

Master's Thesis 2026
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in L^AT_EX
Gothenburg, Sweden 2026

Incorporating AI into the UX Design Process
KIMIA MOHAMMADI MOHAMMADI, ELIN RUUD
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg

Abstract

Artificial Intelligence (AI) tools are becoming a common part of UX design practice and are increasingly being integrated into different stages of the design process. Despite their growing adoption, there is still a limited understanding of how designers experience using these tools, particularly during the early and creative phases of design.

This thesis began with exploratory research to investigate the current use of AI throughout the UX design process and to identify key opportunities and challenges. Based on the findings of this initial study, the research focused on how junior and mid-level UX designers use AI during concept exploration and ideation. To investigate this, two workshop sessions were conducted in which participants used AI tools while completing a concept exploration and ideation task. Data was collected through think-aloud observations, mind maps, love and breakup letters, and focus group discussions. The findings show that designers mainly value AI as a warm-up tool that helps them get started, generate initial directions, and stimulate thinking. However, several challenges were identified, including generic outputs, AI sycophancy, difficulties in creating effective prompts, and the limitations of a linear chat interface for visual and exploratory design work.

Based on these findings, six recommendations were developed and evaluated with the participants to support a more effective and critical use of AI during concept exploration and ideation. The recommendations address prompting strategies, maintaining critical thinking, managing generic outputs, and structuring the ideation process when collaborating with AI.

Keywords: AI in UX design, concept exploration, ideation, design recommendations, human-AI interaction, LLMs, generative AI

Acknowledgements

We would like to express our sincere gratitude to our supervisor, **Martina De Cet**, for her continuous support, guidance, and valuable feedback throughout this thesis. Her encouragement, insightful comments, and dedication have played an important role in shaping both the direction and quality of our work.

We would also like to thank our co-supervisor, **Dr.Mohammad Obaid**, for his scientific expertise, insightful perspectives, and thoughtful advice throughout this research journey. His guidance and support have been invaluable and greatly contributed to the development of this thesis.

Kimia Mohammadi Mohammadi, Elin Ruud, Gothenburg, 2026-06-17

Contents

List of Figures	xiii
List of Tables	xv
1 Introduction	1
1.1 Research Questions	2
1.2 Scope and Delimitations	2
1.3 Thesis Structure	2
2 Background and Related Works	5
2.1 AI in UX Design Process	5
2.1.1 AI in Concept Exploration and Ideation	7
2.1.2 The Risk of Narrowing the Creative Solution Space	8
2.1.3 Purpose-Built AI Tools for Collaborative Ideation	8
2.2 Designer Attitudes Toward AI in Design	9
2.3 AI Adoption in Teams and Organisations	10
3 Theory	13
3.1 User Experience Design	13
3.1.1 The UX Designer's Role and Responsibilities	13
3.1.2 Wicked Problems	14
3.1.3 User-Centered Design	15
3.1.4 Concept Exploration and Ideation in the Design Process	16
3.2 Understanding Artificial Intelligence	16
3.2.1 What is AI?	16
3.2.2 Human and Artificial Intelligence	17
3.2.3 Generative AI	18
4 Methods	21
4.1 Double Diamond	21
4.2 Discover	21
4.2.1 Background Research and Literature Review	22
4.2.2 Survey	22
4.3 Define	22
4.3.1 Quantitative Data Analysis	23
4.3.2 Thematic Analysis	23

4.3.3	Affinity Mapping	23
4.4	Develop	23
4.4.1	Workshop	23
4.4.1.1	Pilot Study	24
4.4.1.2	Think Aloud	24
4.4.1.3	Mind Mapping	24
4.4.1.4	Love and Breakup Letters	25
4.4.1.5	Focus Group	25
4.4.2	Design Recommendations as Prototypes	25
4.5	Deliver	26
4.5.1	Evaluation focus group	26
5	Execution and Process	27
5.1	Initial Survey	27
5.1.1	Participants	27
5.1.2	Structure and Procedure	28
5.1.3	Quantitative Data	29
5.1.3.1	Initial analysis	29
5.1.3.2	Creating graphs	30
5.1.4	Qualitative Data	30
5.1.4.1	Familiarizing with Data	31
5.1.4.2	Initial Codes	31
5.1.4.3	Searching for Themes	31
5.1.4.4	Review Potential Themes	32
5.1.4.5	Defining and Naming Themes	33
5.1.4.6	Write-up	33
5.1.5	Survey Insights Informing the Workshop	38
5.2	Workshop	40
5.2.1	Pilot Study	40
5.2.2	Participants	41
5.2.3	Structure and Procedure	41
5.2.3.1	Part 1: Concept Exploration and Ideation Task	42
5.2.3.2	Part 2: Reflection Activity	43
5.2.3.3	Part 3: Focus Group Discussion	44
5.2.4	Data Collection and Analysis	45
5.2.4.1	Transcript Familiarisation and Coding	45
5.2.4.2	Workshop Workflow Mapping	46
5.2.4.3	Theme Development	46
5.2.4.4	Workshop Themes	48
5.2.4.5	Development of Initial Recommendations	57
5.3	Focus Group	61
5.3.1	Participants	61
5.3.2	Structure and Procedure	61
5.3.3	Data Analysis	62
5.3.4	Final recommendations	62
6	Results	65

6.1	What are UX designers' current experiences of using AI tools across the UX design process? (RQ1)	65
6.1.1	Quantitative Data	65
6.1.2	Result based on thematic analysis	67
6.2	What challenges and opportunities emerge when UX designers interact with AI in concept exploration and ideation? (RQ2)	68
6.3	What are recommendations that can improve the use of AI in concept exploration and ideation, to support UX designers and improve their experience? (RQ3)	69
6.3.1	Recommendation Knowledge	70
6.3.2	Recommendation 1: Direct the output through context	70
6.3.3	Recommendation 2: Challenge and stay critical of AI responses	71
6.3.4	Recommendation 3: Change the settings to customise responses and memory	71
6.3.5	Recommendation 4: Use AI to spark inspiration while keeping creative control	72
6.3.6	Recommendation 5: Generate diverse and rough AI ideas, then explore independently	72
6.3.7	Recommendation 6: Do not use AI for advice, use it to challenge your ideas	73
7	Discussion	75
7.1	Limitations of the Final Recommendations	75
7.2	Risks of using AI in design	76
7.2.1	Risk of deceptive marketing	76
7.2.1.1	Ending or saving the world?	76
7.2.1.2	The design of the tool	77
7.2.2	Impact on Designers' mental health and quality of product	78
7.3	Ethical considerations	79
7.3.1	Bias, cultural values and stereotypes	79
7.3.2	Data labellers	79
7.3.3	Data centres	80
7.3.4	Crediting artists	80
8	Conclusion	81
	Bibliography	83
A	Appendix 1	I
A.1	Mind Maps	I
A.2	Love and break-up Letters	VI
B	Appendix 2	XI
B.1	Introduction to Final Recommendations	XI
B.2	Recommended Knowledge (RK) about AI	XI
B.2.1	RK1: How AI mimics speech	XII
B.2.2	RK2: AI Impacts on designers	XV

B.2.3	RK3 Output validity and Efficiency of Context	XVII
B.3	General AI recommendations across the Ideation Phase	XVIII
B.3.1	Recommendation 1: Direct the output through context	XVIII
B.3.2	Recommendation 2: Challenge and stay critical of AI responses	XIX
B.3.3	Recommendation 3: Change the settings to customise re- sponses and memory	XX
B.4	AI recommendations for specific stages of the ideation phase	XXI
B.4.1	Recommendation 4: Use AI to Spark Inspiration While Keep- ing Creative Control	XXII
B.4.2	Recommendation 5 - Generate Diverse and Rough AI Ideas, Then Explore Independently	XXIII
B.4.3	Recommendation 6 Do not Use AI for Advice, Use It to Chal- lenge Your Ideas	XXIV

List of Figures

4.1	Double design process in this study	22
5.1	Image of the codes and themes laid out on a table when being analysed	32
5.2	Four mind maps from the workshop	43
5.3	Four of the participants' love and break-up letters	44
5.4	Workshop Workflow Mapping. This map was created as a by-product of structuring and analysing the participants' workflow and reflections.	47
5.5	Workshop thematic map	48
6.1	Column chart showing UX designers' level of AI integration in the design process	66
6.2	Column chart showing how useful designers find using AI tools in six different design tasks	66
6.3	Column chart showing how challenging designers find using AI tools in six different design tasks	67
6.4	Chart illustrating the combined results of how useful and challenging designers found AI tools across six design tasks. The red bars indicate the overlap in which designers found AI tools both useful and challenging.	67
A.1	Mind map: Participant 1	I
A.2	Mind map: Participant 2	I
A.3	Mind map: Participant 3	II
A.4	Mind map: Participant 4	II
A.5	Mind map: Participant 5	II
A.6	Mind map: Participant 6	III
A.7	Mind map: Participant 7	III
A.8	Mind map: Participant 8	IV
A.9	Mind map: Participant 9	IV
A.10	Mind map: Participant 10	V
A.11	Mind map: Participant 11	V
B.1	Person speaking a new language fluently, but without knowing what is being said.	XII
B.2	Sketch of what how interaction with Eliza might have looked like. Based on one example from the report [91].	XII

B.3	A simplified sketch of how reinforcement learning changes the weights depending on whether the guess is correct or not. This results in a model that can structure sentences based on which word is most likely to appear next.	XIII
B.4	A sketch of how responses from an AI chatbot might change before and after adapting the model's weights based on what is appropriate to respond and not. This is done through Reinforcement Learning from Human Feedback (RLHF), a training method that uses human preferences to guide the model toward what it should respond. . . .	XIV
B.5	A satirical sketch of how the user is in charge of making the final decisions, but only within the scope of what is being suggested. . . .	XVI

List of Tables

5.1	Overview of workshop participants	42
5.2	Overview of workshop participants, including their group, gender, professional title, and years of experience.	42
5.3	Design insights	57
5.4	Design insights	63

1

Introduction

Artificial intelligence (AI) and large language models are becoming an increasingly common part of people's lives, helping them with different daily tasks. These tools are also becoming an integrated and common part of more professional and technical practices, such as design. Designers now use AI-supported tools to generate ideas, create layouts, produce interface components, summarise user feedback, and support design decisions throughout the design process [1]. These tools can help designers work faster and explore more alternatives. At the same time, AI also introduces new challenges. Designers must decide how much to trust AI-generated suggestions, how to interpret them, how to stay in control of design decisions, and how to maintain their creativity and responsibility while benefiting from these tools [2], [3].

AI can influence design work at multiple stages and in different ways. How AI is used across different phases of the design process is not uniform: its value and impact vary depending on the stage, the task, and the experience level of the designer [4], [5]. In early stages, such as ideation and concept generation, AI supports brainstorming and visual exploration, while in mid-stages it assists with prototyping and iteration, and in later stages it can support evaluation and testing. While the benefits of AI are clear in some areas, it also introduces challenges. These include over-reliance on generated outputs, loss of creative agency, and difficulties in maintaining critical judgment. As a result, some stages of the design process remain more complex and less well understood when it comes to effective AI integration.

Yet despite growing adoption of AI tools in design practice, there is still a limited understanding of how professional designers actually experience using AI across their workflow. Most existing research either examines AI in design broadly without paying enough attention to specific phases such as concept exploration and ideation, or focuses more on introducing new tools, rather than producing meaningful insights and recommendations that can better support designers in practice [5], [6]. Additionally, junior and mid-level designers, who are still developing their professional skills and judgment, remain largely under-represented in existing research, despite being the group that arguably needs the most guidance when working with AI [7], [8].

In the early phases of design, such as concept exploration and ideation, a designer's creativity, reflection, and judgment play a central role. These phases involve exploring a wide range of ideas, making sense of design problems, and deciding which directions to pursue. When AI is introduced into these phases, it can shape how designers think and what solutions they consider. Although AI can inspire new ideas and

support exploration, it can also lead to over-reliance, reduced reflection, or a loss of confidence in one's own reasoning.

Therefore, this thesis addresses this problem by researching and developing recommendations to better support junior- and mid-level designers in the concept exploration and ideation stage of the UX design process. This phase was identified as the main area of focus through an initial investigation that gathered an overview of how UX designers currently use AI across the design process via an online survey. Based on this exploratory research, the focus narrowed to the use of AI for concept exploration and ideation among junior- to mid-level designers, to identify challenges and opportunities, and to provide recommendations that can better support them in this area.

1.1 Research Questions

This study is guided by the following three research questions:

- **RQ1:** What are UX designers' current experiences of using AI tools across the UX design process?
- **RQ2:** What challenges and opportunities emerge when UX designers interact with AI in a specific stage (concept exploration and ideation) of the design process?
- **RQ3:** What are recommendations that can improve the use of AI in a specific stage (concept exploration and ideation) to support UX designers and improve their experience?

1.2 Scope and Delimitations

This thesis focuses on professional UX designers with junior to mid-level experience, examining how they use and experience AI tools in their design work. Exploring how senior designers use AI in their design process warrants investigation, but won't be the focus of this study.

Additionally, the study does not aim to evaluate or compare specific AI tools, nor does it seek to make generalizable claims about designers across all contexts. The findings are exploratory and qualitative in nature, intended to deepen understanding rather than produce universal guidelines. Additionally, the study is conducted in a research context and does not target any specific industry, company, or product type.

1.3 Thesis Structure

This thesis is organised as follows: Chapter 2 presents the background and related work, covering the theoretical foundations of the design process, concept exploration and ideation, and a review of existing research on AI in UX design practice. Chapter

3 presents the theoretical framework, covering key concepts in user experience design, human-AI interaction, and the role of AI in creative and design thinking. Chapter 4 describes the research methodology, which follows a double diamond structure and explains the methods used in each phase. Chapter 5 describes how these methods were applied in this study, including the specific decisions made, the participants involved, and how each part of the study was carried out in practice. Chapter 6 answers the research questions and presents key findings from the study. In the Next chapter 7, the findings and methods are discussed, as well as important factors that were not considered and should be further investigated. Finally, chapter 8 answers the research questions and concludes the thesis with reflections on the study's contributions, limitations, and directions for future research.

2

Background and Related Works

This chapter provides the background and context for the study. Reviewing the existing literature allowed us to familiarise ourselves with research already conducted in the area of AI in UX design, identify key gaps in current knowledge, refine the research focus, and strengthen the recommendations that are the main outcome of this study. It begins by reviewing existing research on how AI is currently used in the UX design process, its role in concept exploration and ideation, and the risks it poses to creative thinking. It then examines purpose-built AI tools that have been developed to better support designers in early ideation. It also covers designer attitudes toward AI and its adoption in teams and organisations.

2.1 AI in UX Design Process

Over the past few years, generative AI has moved from being a peripheral curiosity to an active presence in how many designers work. Large language models such as ChatGPT, Claude, and Gemini are now used across a range of design tasks, while image generation tools like Midjourney and DALL-E are employed for visual exploration. Moreover, Design-specific integrations, such as Figma Make and Figma AI Plugins, bring generative capabilities directly into the tools designers already use on a daily basis [7], [9].

However, despite this rapid growth of AI tools in design, not all areas of design practice have benefited equally. Current AI tools are distributed unevenly across the Double Diamond phases, with some areas, such as wireframing, prototyping, and visual production, receiving more tool support than others. Importantly, [5] highlights that the empathy and cross-screen experiential knowledge that UX designers bring to their work are still significantly underserved by existing AI capabilities. This points to fundamental gaps in how AI currently supports design practice. In particular, the more human dimensions of the work, such as empathy and contextual judgment, remain largely out of reach for current tools.

Another study [4] has examined the usage of AI in the across the entire process and in specific parts of UX design. This study examined how interaction designers use LLMs across all four phases of the design process, and found that AI is very integrated into ideation and early prototyping, while evaluation and user research still require substantial human judgment. The study shows that the arrival of AI in

design practice is not simply a productivity story, but something more structurally complex.

Other studies, such as Tholander et al. [10], focused on a specific part of UX design process. In their workshop study, they focused on ideation and early prototyping, investigating how generative AI and large language models (specifically GPT-3 and DALL-E 2) can play a role in the creative design phases of ideation, early prototyping, and sketching, and how designers conceptualise and relate to AI as a co-creative partner.

In another study, Takaffoli et al. [7] compared different human-AI collaboration workflows for generating personas from survey data. They aimed to understand which workflow, with varying levels of AI involvement, produces the highest-quality personas as judged by designers. The work compared fully manual persona creation, LLM-only generation, and several hybrid approaches where human designers and AI contributed at different stages. The key finding was that LLM-summarising workflows produced personas rated as more consistent, useful, and empathy-evoking than those created by designers alone. However, the study also highlighted that human designers' ability to recognise important user traits augmented the LLM's summarisation in ways that pure AI generation could not replicate. The highest-quality outcomes emerged from workflows where human and AI contributions were genuinely combined, not simply where one replaced the other.

Other researchers have explored the usage of AI for User research and analysis. The question of using LLMs as stand-ins for human participants in qualitative research was examined more critically in a study by Kapania et al [11], where 19 experienced HCI researchers were asked to recreate LLM-simulated versions of their own previous interview participants using GPT-4. While researchers were initially surprised by how plausible the AI responses seemed, they soon identified six fundamental limitations, including the AI's lack of embodied experience, its ambiguous epistemic position, and the risk of delegitimising qualitative ways of knowing. This finding is critical: LLMs may superficially mimic human responses without capturing the lived, contextual knowledge that qualitative research depends on.

Two further studies [12], [13] explored how AI can support the analysis of qualitative data rather than its collection. Qualitative researchers reported using LLMs for initial coding and summarisation with cautious optimism, but raised strong concerns about data privacy when uploading participant data to commercial platforms, and about AI systematically excluding under-represented voices from analysis. A parallel study of 15 HCI researchers found that while they were comfortable with AI assistance in the exploratory stages of qualitative data analysis, they firmly resisted full automation of interpretation, arguing that AI misses subtle nuances that require human contextual judgment. Taken together, these studies draw a clear line: AI is a useful assistant for lower-level user research tasks, but deep interpretive work must remain human-led.

In another study [14], researchers examined how conversational AI assistants could improve UX evaluation tasks, comparing proactive suggestions, offered without being asked, with reactive ones offered only on request. Proactive suggestions helped analysts surface usability issues more efficiently, but over-reliance on AI suggestions

was identified as a notable risk.

Overall, these studies indicate that AI tools are now widely used in design, but their value is uneven. They work best in simpler stages, while more critical tasks requiring deeper human judgment still depend on the designer. The best outcomes come from combining human and AI contributions, and how designers engage with AI matters as much as which tool they use.

2.1.1 AI in Concept Exploration and Ideation

Given the focus of this thesis on concept exploration and ideation, the following section reviews literature that examines how AI supports and limits designers in this specific stage of the design process.

The potential of AI to support early-stage creative work in design is well established in recent literature. Designers consistently report that AI tools are most useful at the beginning of a project, when they need to generate a wide range of directions quickly, break out of habitual thinking patterns, or simply produce something to react to when starting from a blank page [7], [9], [15]. This "getting started" function is one of the most consistently reported benefits across studies, regardless of the specific AI tool being used. However, the picture becomes more nuanced when looking at what AI actually contributes beyond this initial spark.

A study comparing ideation processes with and without generative AI tools found that designers perceived AI-assisted brainstorming as more efficient, particularly for rapidly expanding the initial idea space [15]. Key themes included the importance of accurate verbal prompts, the role of AI as a collaborative thinking partner, and the ability of AI to accelerate the early creative process [15]. However, the study was conducted primarily with design academics and students, and participants raised concerns about the unpredictability of how AI capabilities would evolve, making it difficult to plan long-term workflows around specific tools [15]. This suggests that while AI can accelerate early ideation, designers' ability to benefit from it depends heavily on how they engage with it.

A workshop study that is particularly relevant to this research asked professional designers and HCI researchers to use GPT-3 for design ideation on a real scenario [10]. The workshop involved a half-day session structured around rapid concept development, design reflection, and speculative thinking about future uses of AI in design. The results revealed three main themes. First, AI was most useful for rapidly mapping out a design space and identifying obvious ideas, but participants did not trust it to generate high-quality or innovative solutions independently [10]. Second, AI's lack of contextual memory was a significant practical limitation: it could not track the evolving logic of a design session, meaning designers constantly had to re-establish context [10]. Third, participants described interacting with AI through four distinct framings: conversational, command-based, search-based, and chance-based [10]. The authors argue that moving beyond the conversational metaphor opens up more productive ways of working with AI in creative contexts, and they propose a "designing with" framing in which AI and designers are understood as

doing genuinely different things, rather than competing over the same creative work [10].

Taken together, these findings suggest that AI's value in early ideation is real but narrow: it helps designers get started and react, but it does not replace the judgment and direction that designers themselves must provide.

2.1.2 The Risk of Narrowing the Creative Solution Space

While AI's role as an idea generator is broadly appreciated, a more critical line of research has examined whether AI actually narrows creative thinking rather than expanding it. This is perhaps the most consequential open question for the design profession [6].

A study with 60 participants used a between-participants experiment to test whether exposure to AI-generated images during ideation helped or harmed creative output, measured across design fixation, fluency, variety, and originality [6]. The results were striking: participants who used an AI image generator during ideation produced fewer ideas with less variety and lower originality compared to those who used no tool or used image search alone [6]. The study also identified a phenomenon the authors call "fixation displacement", in which designers chose to focus on AI's visual outputs rather than their own initial examples. The effect on creative output depended heavily on how individual participants approached prompting, suggesting that the way a designer engages with AI is as important as which tool they use. [6].

This finding aligns with what designers themselves reported in qualitative studies. Multiple participants across studies [7], [10] described AI's ideation outputs as generic, predictable, or biased toward mainstream solutions. One key dynamic identified in the literature is that AI systems trained on large datasets tend to produce outputs that reflect the statistical centre of those datasets, meaning they naturally gravitate toward the most common and the most average. For design work that depends on originality and the questioning of conventions, this is a structural limitation rather than a correctable bug [10]. The concern is not only that individual designers may be limited by this tendency, but that if many designers across the profession are using the same AI tools for early concept generation, the collective creative output of the field may become more homogeneous over time.

In short, AI may help designers generate more ideas, but not necessarily better or more original ones. When designers follow AI's lead too closely, they risk narrowing rather than expanding their creative thinking, a concern that goes beyond individual designers to the profession as a whole.

2.1.3 Purpose-Built AI Tools for Collaborative Ideation

Recent work has begun to address the mismatch between what designers need in early ideation and what general-purpose AI tools currently offer. AIdeation is a purpose-built human-AI ideation system designed specifically for concept designers, addressing challenges around finding relevant references, working under time pressure,

and generating diverse directions [16]. Unlike general image generation tools such as Midjourney or DALL-E, which were not designed for the reference-driven and iterative workflows of concept designers, AIdeation supports staged exploration that aligns more closely with how concept designers actually work [16]. Its evaluation found that designers using it could find references more aligned with their creative vision and explore more diverse directions within constrained time conditions [16].

IdeationWeb takes a different approach by focusing not on generating ideas but on visualising and tracking how ideas evolve during human-AI co-creative sessions [17]. The system helped designers maintain awareness of how AI was influencing their ideation paths, surfacing patterns of over-reliance on certain AI-generated directions, and supporting more intentional and reflective co-creative practice [17]. The ability to see one’s own creative trajectory, including where it diverged from or converged with AI suggestions, gave designers a new kind of metacognitive resource during ideation [17].

These tool-level developments reflect a broader shift in how researchers are framing the problem. Rather than asking whether AI can replace designers in ideation, the more productive question is how AI tools can be designed to augment and diversify creative thinking rather than anchor it [16], [17]. Overall, the most promising direction is not building AI that generates more ideas, but building AI that helps designers think more intentionally, tools that make the creative process more visible, more diverse, and more reflective.

2.2 Designer Attitudes Toward AI in Design

Designers generally see AI as a useful assistant in their work rather than a replacement for their own thinking. A survey of 51 UX professionals found that practitioners value AI most as a creative partner during iteration and for taking over routine, time-consuming tasks, freeing them to focus on the parts of the design process that require genuine human judgment [18]. This positive framing is consistent: designers do not feel threatened by AI, but they do draw clear lines around what they are and are not comfortable delegating to it.

At the same time, over-reliance stands out as one of the most commonly raised concerns. When AI is always available and quick to produce a plausible-looking output, it becomes easy for designers to accept those outputs without applying their own critical thinking [18]. This is particularly relevant during creative tasks, where the quality of a designer’s judgment is what separates a genuinely good solution from one that merely looks adequate on the surface.

The risk is not that AI produces obviously bad work, but that it produces work that is good enough to stop designers from pushing further. Research on over-reliance on AI shows that users tend to accept plausible-looking outputs without enough questioning, particularly after a few positive experiences with a tool [19]. In creative professions especially, this is compounded by cognitive offloading, where designers increasingly delegate thinking to AI. By doing so, they risk engaging with design

problems only on the surface and accepting the first reasonable answer rather than exploring deeper, more considered solutions [20], [21].

Beyond over-reliance, designers also carry broader concerns about what AI means for their practice as a whole. Survey data from UI/UX professionals identifies algorithmic bias in AI-generated outputs as a primary worry, since AI tools trained on existing data can reflect and reinforce patterns that do not serve all users equally [22]. Data privacy is another consistent concern, where uploading sensitive user research or client project information to commercial AI services creates risks that many designers feel are not yet adequately addressed by the tools or their organisations [22].

Taken together, the picture that emerges is one of cautious optimism. Designers welcome what AI can do for them and have found real value in working with it. But they are clear-eyed about the risks, and their concerns point to something important: the value of AI in design depends not just on the quality of its outputs, but on the conditions under which designers engage with it [18], [22].

2.3 AI Adoption in Teams and Organisations

While individual designers have been quick to explore AI tools, the picture at the team and organisational level looks quite different. Research across 24 UX practitioners in industry found that most companies have no clear policies around AI use in design, and teams lack coordinated practices for deciding when, how, or even whether to use these tools [7]. In practice, this means AI adoption is happening informally and inconsistently: one designer on a team may rely heavily on AI for certain tasks while a colleague avoids it entirely, with no shared framework to align their approaches or evaluate the results [7].

This lack of organisational structure creates real friction. In teams where designers and engineers work together on AI-augmented products, research found that UX designers often lack the technical fluency to guide AI model behaviour directly [23]. Instead, they rely on engineers to translate their design intent into effective prompts or technical configurations, creating a new kind of collaboration gap that existing team structures were not built to handle [23]. The result is that the benefit designers get from AI tools depends partly on who they happen to work with, rather than on their own skills or the quality of the tool itself.

Context also plays a significant role in how useful AI actually is within a company setting. Designers working on a single, well-defined product in a stable organisational context tend to get more out of AI than those working at agencies, where the client and project change constantly [7].

Research on the application of AI in controlled organisational settings presents a nuanced picture. A study involving eight senior UX design experts found that using ChatGPT for creating design briefs improved both efficiency and the overall task experience. This suggests that AI can provide significant value within structured workflows. However, challenges persist, particularly in terms of protecting content copyright and establishing professional identities. One of the most crucial factors

remains the assurance of information authenticity [8].

Overall, these studies show that integrating AI into design teams is not just a matter of individual skill. It requires organisational support, shared practices, and clearer frameworks, without which AI adoption remains inconsistent and its benefits hard to realise.

3

Theory

This chapter presents the theoretical foundations that support the study. It is organised into two main sections. The first covers key concepts in user experience design, including the role of the UX designer, the nature of wicked problems, user-centred design, the Double Diamond model, and the place of concept exploration and ideation in the design process. The second section focuses on human-AI interaction, examining what AI is, how generative AI works, and how humans perceive and relate to AI systems. These theoretical perspectives provide the conceptual grounding for this research.

3.1 User Experience Design

User Experience Design (UXD) focuses on creating meaningful and clear experiences for users when interacting with products, services, or systems. The primary goal of an appropriate user experience is to meet the real needs of users. Although this goal sometimes seems easy to reach, the process of achieving it might get complicated [24]. Although the process of UXD has changed over the past decades, it primarily involves understanding user needs and wants, developing design strategies, prototyping, and evaluating the usability of a product, system, or service to ensure a seamless and engaging user journey. UXD also plays a critical role in fostering emotional connections between users and products, contributing to brand loyalty and satisfaction. In today's technological landscape, the integration of AI into the design process has not only changed the UX design process but is also increasingly shaping how UX designers work. Understanding UXD as a discipline is therefore essential to this study. In fact, without a clear picture of what UX design entails and what it aims to achieve, it is difficult to understand how and where AI integration has an impact. This also makes it important to first understand who the UX designer is, what challenges they face, and how their process is structured [25], as these are the exact aspects this thesis investigates in relation to AI. The following subsections explore each of these in more detail.

3.1.1 The UX Designer's Role and Responsibilities

Understanding what UX designers do and are responsible for is essential for this study, as it clarifies where in the process AI may offer support and what tasks require human judgment and expertise. For example, UX designers are fundamentally responsible for

creating user-centred, intuitive, and efficient digital experiences. Their responsibilities span across various stages of the design process and often involve collaboration with multidisciplinary teams. Some of these key responsibilities are as follows:

- **User Research and Data Synthesis:** UX designers are responsible for conducting user studies to understand users’ needs, preferences, and pain points. They gather both qualitative and quantitative data, and also synthesise insights to inform design decisions [26].
- **Ideation and Solution:** After defining the problem, users’ needs, pain points, and opportunities, UX designers use various methods to ideate and develop solutions. In this process, they should be creative while ensuring that user needs remain central. Integrating technologies such as AI tools may support this process, which is the focus of this thesis [27], [28].
- **Design and prototyping:** Some designers are responsible for creating prototypes, wireframes, and user journeys, although these tasks may also be handled by other roles depending on the design field. These artefacts help visualise and test design concepts before full-scale development. Designers often work iteratively, refining their work based on user feedback and usability testing [29].
- **Evaluation and testing:** Evaluation is another important part of their responsibilities, as it is essential to improve the product. In fact, prototypes are tested with real users to uncover friction, confusion, or failure points. Multiple rounds of testing typically occur, with each round informing further refinement [30].

Apart from the responsibilities outlined above, UX designers are also expected to work closely with multidisciplinary teams, including developers, marketers, and other designers, to align design goals with broader business objectives. They also bridge the gap between UX terminology and organisational language to ensure that stakeholders understand the value of UX work [31]. Additionally, they are responsible for ethical and sustainable design. This includes incorporating sustainability into their practice and creating scalable, future-proof products that remain responsible and relevant as user needs and technologies continue to evolve [32].

3.1.2 Wicked Problems

Design problems are rarely straightforward. Understanding why helps clarify both the nature of the designer’s work and the limitations of AI in supporting it. The concept of wicked problems explains why design requires iterative, reflective thinking, and why AI tools alone are not enough to solve them.

In 1973, Rittel and Webber [33] introduced the concept of “wicked problems” to distinguish complex and unstructured problems from what they described as “tame” problems. Tame problems are well-defined, have clear boundaries and goals, and can be addressed using established scientific or technical methods that lead to verifiable solutions. These types of problems have been successfully solved using AI systems [34]. In contrast, wicked problems are characterised by ambiguous goals,

lack a clear end, and have no true or false solution, only better and worse [33]. Additionally, each attempted solution changes the nature of the problem itself and is therefore a ‘one-shot operation’ that requires careful examination before it is executed. Therefore, Rittel and Webber argued that such problems cannot be solved with traditional problem-solving methods used in science. Instead, they proposed that wicked problems require iterative, reflective, and context-sensitive strategies that emphasise learning, reframing, and stakeholder negotiation. This process requires the ability to reason and reflect, which current AI systems lack. Many contemporary design processes, including the double diamond, continue to reflect these principles by emphasising exploration, prototyping, and continuous evaluation.

These characteristics are not unique to one design discipline. Although design encompasses many specialised domains, such as graphic design, industrial design, and interaction design, they all share the same fundamental challenge of understanding and addressing wicked problems. As a result, designers across these fields rely on similar iterative and reflective processes that emphasise exploration, experimentation, and continuous refinement. This also extends to the relationship between design and AI. There is a clear parallel between the characteristics of wicked problems and the current abilities of AI systems, particularly in their difficulty handling ambiguity, contextual interpretation, and value-based judgment. This connection highlights why the concept of wicked problems remains highly relevant for understanding the role and boundaries of AI in UX design.

3.1.3 User-Centered Design

User-Centred Design (UCD) is a design approach that prioritises the needs, goals, and experiences of users rather than focusing solely on the product or technology. This is the core philosophy behind how UX designers approach their work. It matters for this study because it defines the human values and practices crucial to the design process, and these values should therefore be considered when investigating how AI tools can support the design process.

User-Centred Design includes three main principles: the involvement of users throughout the entire design process, iterative design cycles, and ensuring that the final product aligns with user expectations [35]. The first core principle of User-Centred Design is user involvement, meaning that users are actively engaged at every stage of the design process to understand their needs, tasks, and goals. Another core principle is the iterative process: the design goes through multiple cycles, where prototypes are tested and refined based on user feedback to improve usability and user experience (UX). The third principle is a multidisciplinary approach: UCD brings together people with different areas of expertise to ensure that diverse perspectives inform design decisions. Finally, UCD aims to create products that are efficient, easy to use, and enjoyable, addressing both usability and experiential quality [36].

3.1.4 Concept Exploration and Ideation in the Design Process

Concept exploration and ideation are the specific stage of the design process that this thesis examines. Understanding what this stage involves and why they are important forms the foundation for the research questions and the workshop study in this thesis.

Concept exploration and ideation are important parts of the UXD process. Concept exploration involves broadly investigating possible directions for a solution. During this phase, designers may explore different product ideas, features, interaction models, or even entirely different approaches to solving the problem [37], [38]. This phase is inherently divergent, encouraging designers to suspend judgment and remain open to a wide range of possibilities before narrowing down to a specific direction [39].

Ideation, on the other hand, refers to a structured creative process of generating a wide range of ideas within these explored directions and finding novel solutions around a set of requirements [40]. It is closely linked to creative thinking and problem reframing, as designers must move beyond obvious solutions to produce ideas that are both novel and relevant to user needs [38]. The goal of ideation is not to find the perfect idea immediately, but to generate enough variety that promising directions can be identified, combined, and refined in later stages of the design process.

3.2 Understanding Artificial Intelligence

AI is at the centre of this study, and understanding what it is, how it works, and how it differs from human intelligence is essential for evaluating its role in the UX design process. This section provides that foundation. This section provides the theoretical background on artificial intelligence relevant to this study. It begins by clarifying what AI is, then examines how it differs from human intelligence, and finally explains how generative AI works.

3.2.1 What is AI?

The use of AI has increased rapidly since the end of 2022, when ChatGPT, an AI chatbot based on a large language model (LLM), was launched. However, AI has been applied in many fields long before this development, often under different names, such as calculators, chess computers, and medical diagnostic systems [34]. These systems are examples of narrow AI, designed to perform specific tasks within limited domains. This type of AI can achieve high levels of accuracy, reliability, and performance within their specialized contexts. In contrast, general AI, which would be capable of solving a wide range of problems with human-like flexibility, remains largely theoretical. What is commonly referred to as AI today is primarily generative artificial intelligence, which generates text, images, and other forms of content and has been developed and researched for many years prior to its public breakthrough. Understanding what AI is and what it is not is important for addressing common misconceptions that could affect how it is being used. It is especially important

for this study, as misconceptions about AI's capabilities could affect how designers engage with it in practice, and how its integration into the design process is evaluated.

3.2.2 Human and Artificial Intelligence

Scientific evidence strongly suggests that contemporary AI systems do not possess consciousness or the ability to think in a human sense [41]. However, the tendency to attribute consciousness to AI systems is equally important to understand, as it is not merely the result of insufficient knowledge but can be explained through insights from cognitive science. This matters for this study because how designers perceive AI directly influences how they use it, and whether they over-rely on it or engage with it critically. Therefore, understanding how we perceive AI systems, as well as the differences between human and artificial intelligence, is essential to understanding its role in the design process. To understand why this misperception occurs, it helps to first look at how human intelligence and consciousness actually work.

Human intelligence has long been compared to computer processing. In early cognitive science, when the brain was still largely a mystery, and computers started showing signs of human intelligence, comparisons between the two began to emerge [41]. This theory has not been successfully confirmed, despite many years of attempts to identify structures in the brain corresponding to a centralised processor and separate memory storage, as found in traditional computer architectures. Since the 1970s, the theories have instead shifted towards viewing the brain as a neural network, an idea that has inspired the structure of current AI systems. This perspective is still considered relevant when studying the brain. However, when comparing AI models to the human brain as evidence of intelligence, Gärdenfors emphasises that it is important to recognise that the brain is far more complex and consists of more than just a neural network alone. Throughout the book, Gärdenfors gives several examples of how human intelligence differs immensely from artificial intelligence. An example of this is the lack of neurotransmitters such as dopamine and adrenaline. Another is the importance of bodily functions that perceive information and help us create contextual understanding. Consequently, Gärdenfors suggests that thinking emerges through interaction between the brain and the surrounding environment. Liu et al. [42] and McMeel et al. [43] support this perspective in their examination of what would be required to achieve true artificial intelligence. They expand on the concept of logic, arguing that it is highly diverse and is not confined solely to the mind. Instead, logic is shaped by individuals, environments, sensory experiences, and contextual factors, suggesting that intelligence emerges from a broader interaction between cognition and the world. One key aspect of this complexity is conscious experience itself.

Conscious experience, or awareness, can be understood as the information that allows individuals to recognise and reflect on what they attend to and experience [41]. This mechanism not only applies to self-awareness but also shapes how people perceive and interpret others. In everyday interactions, humans constantly attempt to understand what others are thinking and feeling to predict their behaviour and collaborate effectively. These interpretations are based on incomplete information and therefore

require people to fill in gaps using prior knowledge, expectations, and by projecting aspects of their own conscious experience onto others. However, our experience of others is not always accurate. The author uses one of Graziano’s examples [44], which illustrates this tendency through the experience of ventriloquism, in which coordinated mouth movements and body language create the illusion that a puppet possesses awareness and intention. Similarly, human-like language and coherent responses in AI systems can trigger perceptions of intelligence and consciousness, even when these systems lack genuine understanding.

These perspectives highlight fundamental differences between human and artificial intelligence. Unlike humans, AI systems do not have a body, sensory experiences, or direct engagement with the physical world. As a result, they are limited in their ability to perceive our reality, form rich contextual understanding, and ground responses in reasoning and lived experience. This restricts its capacity to resonate with situations, interpret meaning, and apply logic flexibly and in context. Consequently, while AI can model certain cognitive processes, it lacks a crucial foundation for human-like understanding, reinforcing its limitations as a substitute for human intelligence in complex and creative domains such as design.

This reflects a dilemma similar to that identified for wicked problems, where scientific and analytical methods proved insufficient because such problems require reflective and iterative thinking [33]. Narrow AI systems excel at performing specific, well-defined tasks and can significantly accelerate work in structured contexts [34]. However, as discussed earlier, genAI cannot reason, reflect, and develop genuinely original solutions [41], [42], which is necessary to handle wicked problems. Nevertheless, this does not imply that AI cannot be used efficiently in design. Rather, it hints towards which aspects of the design process are suitable for AI support and which require human judgment. Therefore, the following subsection examines how generative AI works and why it lacks the ability to solve wicked problems.

3.2.3 Generative AI

Since this study focuses specifically on how designers use generative AI tools, understanding how these systems work is necessary to interpret their potential and limitations in a design context. GenAI can be communicated with through prompts and creates new content such as text, images, audio, code, and other media. Unlike traditional rule-based or purely predictive systems, genAI systems operate by learning statistical patterns from vast amounts of data [34]. This enables them to produce outputs that resemble human-created content.

An example of generative AI is large language models, such as ChatGPT, Google’s Gemini, and Anthropic’s Claude. These systems generate text by predicting and selecting one word at a time based on the context of the user’s prompt [34], [43], [45]. Each word is chosen from a range of possible alternatives according to its statistical likelihood of fitting coherently within the ongoing text.

These statistical probabilities are created by training the model using large collections of text data. During training, the model learns by repeatedly guessing the last word

in a sentence [34], [45]. At the beginning of this process, the probabilities assigned to different words are set randomly, which means that the models' initial predictions are often incorrect. After each prediction, the model compares its chosen word with the correct word in the training data and adjusts its internal probabilities accordingly. This model is inspired by the structure of the human brain and its neural networks, but is implemented through an artificial neural net consisting of interconnected nodes and weighted connections [41], [45]. The weighted connections hold the probability score that is updated during training to strengthen useful patterns and weaken ineffective ones. As a result, the model learns to generate fluent text based on statistically learned relationships rather than genuine understanding.

4

Methods

This chapter presents the methods used in this study. The research followed the Double Diamond framework, which is described below. For each phase of the Double Diamond, the methods that guided the research are outlined and explained.

4.1 Double Diamond

The Double Diamond is a design process framework introduced by the British Design Council. It not only serves as a design framework but also as a methodological structure, helping to move from a broad understanding toward a focused investigation of a specific stage, and finally to the development of a final product [46], [47]. It consists of four main phases: **Discover**, **Define**, **Develop**, and **Deliver**.

The first phase, **Discover**, focuses on broadening the understanding and exploring the problem space.

This sets the stage for the second phase, **Define**, which involves analysing the data gathered in the Discover phase to identify patterns and define the problem space.

The third phase, **Develop**, is where the main investigative work takes place. Building on the insights from the Define phase, an iteration of ideation and prototyping is being conducted.

The fourth and last phase, **Deliver**, involves testing, evaluating, and, based on this, developing a final product.

The research was carried out in these four phases, and Figure A.11 illustrates an overview of how the study was conducted within the Double Diamond design framework. Below, the methods used within each phase will be described in more detail.

4.2 Discover

During the Discover phase, the focus was on gaining a broad understanding of how UX designers use AI across the design process. This was achieved by reviewing existing literature and current research in this area, as well as conducting an online survey. Each of these activities is described in greater detail in the following subsections.

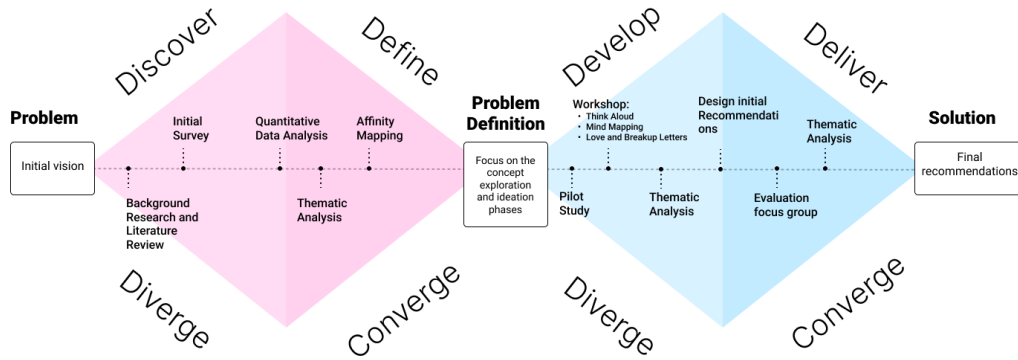


Figure 4.1: Double design process in this study

4.2.1 Background Research and Literature Review

A literature review is a systematic review of existing research and scholarly work relevant to a given topic. It is commonly used in the early stages of a study to map what is already known, identify gaps in existing knowledge, and establish a theoretical foundation for the research [48]. A literature review was conducted at the beginning of this study to understand how AI is currently used in design practice and to establish the theoretical foundation for the survey and workshop.

4.2.2 Survey

User research is a method of collecting data from users and analysing it with the aim of reaching the insights needed for developing a product, service, or system [49]. A survey is a data collection method used to gather information from a group of people through a set of structured questions. It is particularly well-suited for reaching a larger number of participants within a limited time frame and for collecting both quantitative and qualitative data in a single instrument. Surveys allow researchers to identify patterns across a sample and compare responses based on background variables such as role, experience level, or context of use. In this study, the user research began with a survey to collect data about how UX designers use AI tools across different phases of the design process.

4.3 Define

During the Define phase, the data gathered in the Discover phase was analysed to identify patterns and narrow the focus of the study. The methods used to analyse the data gathered from the survey include both quantitative and qualitative approaches, which are explained in more detail in the following subsections.

4.3.1 Quantitative Data Analysis

Quantitative data analysis is a method used to examine numerical data to identify patterns, trends, and relationships within a dataset. It typically involves summarising and representing data through descriptive statistics, visualisations such as graphs and charts, and comparisons between groups [50].

4.3.2 Thematic Analysis

Thematic analysis is a qualitative research method for analysing qualitative data [51]. This method is used to identify, analyse, and interpret patterns of meaning within a dataset. Given its flexibility, this method is well-suited to understanding experiences, attitudes, or behaviours rather than measuring them.

Thematic analysis can be applied either inductively or deductively [51]. In an inductive approach, themes emerge directly from the data without being guided by a pre-existing theory or framework. In a deductive approach, the analysis is guided by predefined categories or research questions.

It is important to provide a transparent result, showing how themes were developed [51], [52]. Ahmed et al. [52] supplement the guidelines from Braun and Clarke [51] by adding recommendations of how to minimise personal bias.

4.3.3 Affinity Mapping

Affinity mapping is a method used to organise large amounts of qualitative data into meaningful patterns and themes [53], [54]. This method is particularly useful during research analysis to help synthesise information from user studies and transform complex, unstructured data into clear and structured insights. The method typically involves writing individual observations, codes, or ideas onto separate notes and then grouping them based on similarities. Overarching themes and categories emerge during this grouping process, revealing patterns that might not be immediately visible in the raw data. Involving multiple researchers in the process adds diverse perspectives and leads to a more comprehensive understanding of the data.

4.4 Develop

After identifying the focus area in the define phase, the develop phase began gathering and generating new insights through a workshop study. The findings from this phase were then used to develop an initial set of recommendations. The workshop study, the methods used during it are described in more detail in the following subsections.

4.4.1 Workshop

Workshop studies are a qualitative research method widely used in design research to observe participants engaging with realistic tasks in a facilitated, collaborative setting. In the context of HCI and UX research, unlike surveys or interviews, workshops

enable researchers to directly observe behaviour as it unfolds. As a result, this method helps in capturing not only what participants say but how they act and make decisions in practice [55].

Workshops are particularly valuable as they let researchers collaborate with participants and collect rich insights through shared discussions and activities. The workshop activities let participants reflect on current situations, exchange perspectives, and explore possible future developments. They also often include hands-on and generative design activities to gain more in-depth knowledge around the explored topic [56]. More details about the methods used during the workshop are described in the following sections.

4.4.1.1 Pilot Study

A pilot study is a small-scale, preliminary version of the main study, designed to test the feasibility of the research methods and procedures before they are applied to the actual participants [57], [58]. By simulating the full study process in advance, a pilot study allows researchers to identify and address potential challenges, such as unclear instructions, timing issues, or problems with the study structure that could otherwise compromise the quality of the main data collection. Conducting a pilot study is therefore an important step in ensuring that the main sessions run smoothly and that participants can focus on the task itself rather than on navigating an unclear or poorly structured process [59], [60]. Prior to the main workshop sessions, a pilot study was conducted to ensure that the structure, tasks, and instructions were clear and that the sessions could run as intended.

4.4.1.2 Think Aloud

Think-aloud is a method in which participants are asked to verbalise their thoughts while doing a task. This method allows observation of how participants approach a problem in real time rather than just relying on their personal narratives. As a result, it makes the reasoning and decision-making process visible to the researchers. This method is commonly used in usability studies and design research to find the reasoning behind decisions that would otherwise remain invisible [61], [62].

4.4.1.3 Mind Mapping

Mind mapping is a visual method for extracting a simplified version of the user's mental model. A mental model is our understanding of how something works. [63]. It is a reasoning mechanism that helps us make sense of the world and decide how we should interact with it.

A mind map is conducted by organising thoughts into a map, with associations drawn towards the middle where the central system is situated [64], [65]. Associations are drawn closer or further from the centre depending on how strong their relationship to the system is. The participant is limited to one side of a paper to provide clear boundaries and make the process less overwhelming.

A mind map was used as the base for participants in one of the workshop activities to express and draw their mental model, but they were not restricted by the original description above. Instead, they were allowed to create more complex, detailed maps showing the relationships between associations and functions of the system. However, they were still limited to one page and had an additional time limit to extract their overarching understanding and avoid overwhelming the participants. The umbrella term *cognitive maps* might have been a more appropriate description[65], but it also increases expectations of what the map should contain and was therefore not referred to.

4.4.1.4 Love and Breakup Letters

The love and breakup letter method is a qualitative research technique used to explore the emotional and relational dynamics between users and products, services, or technologies [64]. Participants are asked to write either a love letter or a breakup letter to a tool or product, expressing their feelings, experiences, and reasons for attachment or detachment [64], [66]. By framing the reflection as a personal relationship, the method encourages participants to articulate what works, what does not, and why, in a way that is both honest and emotionally expressive. This makes it particularly well-suited to capturing nuanced experiential insights that are often difficult to surface through more structured data collection methods such as surveys or closed questionnaires. It encourages emotionally honest reflection and generates rich qualitative data about both the positive and negative dimensions of an experience in a form that is easy to discuss and compare across participants.

4.4.1.5 Focus Group

A focus group is a qualitative research method in which a small group of participants gathers to discuss a specific topic under the guidance of a moderator. The primary characteristic of focus groups is the interaction between participants; rather than collecting individual responses in isolation, the method generates data through group dialogue, where participants respond to, build on, and challenge each other's perspectives [67]. This dynamic makes focus groups particularly well-suited to exploring shared experiences and generating rich, in-depth insights that individual interviews or surveys are less likely to surface.

4.4.2 Design Recommendations as Prototypes

The primary output of a development phase in the interaction design and HCI research field is not always a physical prototype. In some studies, the research goal is to improve practice rather than to create a product. In this case, the research contributions are design recommendations for using AI in concept exploration and ideation, grounded in the findings from the literature review, survey and workshop. Recommendations translate empirical findings and analytical insights into actionable guidance that practitioners can apply in their work.

4.5 Deliver

The deliver phase focused on refining the recommendations built in the previous phase. To this end, the recommendations were evaluated, and based on the results of this evaluation, they were revised into a final set of recommendations.

4.5.1 Evaluation focus group

An evaluation session is a method used in design research to assess and refine a proposed artefact, framework, or set of recommendations with the people who would use them in practice. It provides an opportunity to gather critical feedback before finalising the output, ensuring that the output is grounded in real needs and is practical and useful.

5

Execution and Process

This chapter presents how the methods were used to collect and analyse data for the study. The research combined three main methods: 1) an initial survey (see section 5.1), 2) a workshop (see section 5.2), and 3) a follow-up focus group (see section 5.3). The survey was used to gather broader insight into designers' experiences, attitudes, and use of AI tools. Based on this, we decided to focus on concept exploration and ideation, as these phases showed high AI integration while many of the same designers also found it challenging to use AI in these phases. The workshop was then used to explore these phases of the design process in more depth, to support the creation of recommendations. Finally, the follow-up focus group was conducted to assess and discuss the recommendations developed from the workshop, which were then adjusted accordingly.

Using these three methods supported data triangulation, which is important for strengthening the validity of the collected data [cite Interaction Design: Beyond human-computer interaction]. Data triangulation was achieved by collecting data at different times and in different places, but not with a different set of participants, since workshop participants were recruited from the survey and then also present in the final focus group. Researcher triangulation was also supported by our involvement, the two authors, who both contributed to data collection and analysis. Together, these methods provided a broader and more nuanced foundation for the analysis, as the findings were based on multiple data-gathering methods rather than relying on a single method.

5.1 Initial Survey

The initial survey aimed to identify a specific phase in the design process to focus on. To do this, data were gathered through an online survey on which AI tools UX designers use, in which phases, and on their attitudes towards them. An online survey was chosen as the data collection method because it can reach a large number of participants within a limited time frame.

5.1.1 Participants

The study population was defined as UX designers, aligning with the thesis's focus group. However, due to possible difficulties of finding a sufficient number of par-

ticipants working mainly as UX designers, the study population was broadened to include related design roles such as UX researcher, UI designer, product designer, interaction designer, and service designer, with the option to add additional roles. This broader definition was considered relevant because these roles often overlap with UX design in practice. Including related design roles made it possible to collect more data without moving too far from the study's focus, and it managed to reach 52 designers in total, of whom 24 were UX designers.

Participants were recruited through a combination of convenience sampling and snowball sampling. Convenience sampling was used by posting the survey online and contacting UX designers directly via email. Snowball sampling was also used, with participants asked to share the survey with others who fit the target group. This approach was appropriate because the study required participants with specific professional experience, and personal networks facilitated recruitment of individuals within the design field.

5.1.2 Structure and Procedure

The questionnaire was designed with guidance from *Interaction Design: Beyond Human-Computer Interaction*[68]. The guidelines considered during the design of the questionnaire included informed consent, question order and structure, relevant background questions, and what types of questions to use in different situations.

At the beginning of the questionnaire, participants were informed of the purpose of the study and how their responses would be used. This was included to support informed consent. The introduction also gave clear instructions on how to complete the survey, so that participants understood what was expected before answering.

The order of the questions was considered, starting with general background questions before moving into questions about AI use and attitudes toward AI. This structure was chosen to make the survey easier to follow and to avoid asking more complex or reflective questions too early. The questions were also grouped by topic, facilitating a logical progression through the survey.

The first part of the questionnaire contained background questions relevant to the study. These questions first asked about general background, including the participants' gender, age, highest level of education, and country. They then moved on to more design-related background questions, including design role, company size, experience level, and design domain. This information was necessary to understand the context of each response, assess whether the participant belonged to the intended study population, and identify any related patterns in AI use across participants.

Since the aim was to explore not only whether designers use AI but also when, how, and why they use it, the survey required both closed-ended and open-ended questions. Closed-ended questions, such as multiple-choice and rating-scale questions, were used to collect comparable information on participants' background, AI use, and attitudes. Open-ended questions were included to capture more detailed explanations, examples, and reflections. This was important because the study not only aimed to measure whether designers use AI, but also to understand how and why they use it.

The survey was divided into different paths depending on whether participants used AI tools. This was important because non-AI users would not be able to answer detailed questions about personal AI use. Since they may still have valuable insight about why they do not use AI tools, they were directed to questions that focused more on perceptions, expectations, or reasons for not using AI. This made the questionnaire relevant to both AI users and non-AI users.

Before the questionnaire was distributed, feedback was gathered from the supervisor. The supervisor's input was used to assess whether the questions were appropriate for the study's aim and whether the questionnaire's structure was clear. Based on this feedback, adjustments were made to improve the survey before it was shared with participants. These adjustments concerned aspects such as wording, question order, response alternatives, and the distinction between questions for AI users and non-AI users.

5.1.3 Quantitative Data

The survey responses were collated into spreadsheets using Google Sheets and Numbers. The data was structured with each row representing one participant's response and each column representing one survey question. This enabled comparison of responses across participants and filtering the data by background variables. For questions with multiple options, each possible answer was divided into a separate column. This made the data easier to summarise, count, and filter. For example, if a question asked which AI tools participants used, each tool was given its own column, allowing us to see how many participants selected that option.

5.1.3.1 Initial analysis

The initial analysis began by reviewing the spreadsheet to get an overview of the full dataset. The data was colour-coded to make patterns easier to identify visually. This made repeated answers, high or low values, and notable differences among participants easier to see.

The data was also filtered based on background information. For example, responses from UX designers were compared with those from designers in other related roles, such as UX researchers, UI designers, product designers, interaction designers, and service designers. This enabled an overview of how UX designers answered and whether their responses differed from those of the broader group of design participants. Other filters, such as experience level or AI use, were also considered where relevant.

For questions in which the data could be ordered numerically, averages and medians were calculated. This was done for questions such as trust in AI tools and the level of AI integration in the design process. The mean was used to understand the overall tendency within the sample, whereas the median was used to mitigate the influence of extreme values. Using both measures enabled a more balanced understanding of the participants' responses.

5.1.3.2 Creating graphs

Graphs were created to visualise the quantitative findings and make the patterns in the data easier to interpret. The graphs were created for both the full sample and a filtered selection of UX designers compared with non-UX designers. This separation was important because the research question focused on UX design, while the final participant group also included designers from related fields.

For this reason, some graphs showed the entire sample to provide an overview of general tendencies across all participants. Other graphs focused solely on UX designers or compared UX designers with non-UX designers to examine whether there were differences between these groups. This allowed us to assess whether the broader sample reflected the same tendencies as the UX-specific sample, or whether the UX designers showed distinct patterns.

Scatter plots were used to explore relationships between two ordered variables. For example, they could be used to compare participants' levels of trust in AI with the degree of AI integration in their design process. This type of graph made it easier to see whether participants who reported higher AI integration also tended to report higher trust, or whether the relationship was more varied.

Bar graphs were used for categorical data, such as design roles, AI tools used, purposes for using AI, and selected benefits or concerns. These graphs were suitable because they enabled comparison of the frequency with which different options were selected. They were especially useful for multiple-choice and checkbox questions where the goal was to show the distribution of responses across predefined categories.

The graphs were not used to claim statistically significant relationships, as the sample size and sampling method did not support such a conclusion. Instead, they were used as descriptive tools to support the analysis, identify trends, and clarify comparisons between participant groups.

5.1.4 Qualitative Data

The qualitative result from the survey was analysed using thematic analysis. The thematic analysis used was mainly based on Braun and Clarke's paper [51]. The introduction of their paper highlights how personal bias and assumptions affect the development of themes, which are an important part of each step. However, the description of each step provides little guidance on how to achieve this. Therefore, guidelines from a paper by Ahmed et al. [52], who adapted Braun and Clarke's [51] guidelines to include more detailed steps to keep researchers aware of personal biases and assumptions and to make the final themes transparent, were also used.

The analysis aimed to summarise aspects of the data related to how participants use AI and where they found the tool challenging. This was primarily done through an inductive approach, without seeking specific use cases or challenges. The findings were primarily based on a semantic approach, allowing the data to speak for itself. However, some interpretations were offered regarding the underlying needs and reasons for the challenges.

5.1.4.1 Familiarizing with Data

The open-ended questions were extracted from the survey and carefully read through. It is emphasised that familiarising with the data is a crucial step in identifying meaningful patterns and themes [51]. A recommended first step is to manually transcribe the data. However, because the data were written, a brief summary of each response was provided instead. Similar to transcribing, this activity demands attention to the nuances of the data. On the contrary, the summary primarily served as a starting point for immersing in the data and did not replace the original responses when later analysing them. The summary also serves as a quick reference when glancing over responses to find specific codes and content.

After writing the summary, the responses were read through several times. To ensure engagement, an iterative process of marking recurring words and writing down initial ideas, personal assumptions, and observations was taking place while reading. This is an important step to ensure immersive reading, which is required when familiarising with the data [51]. As recommended, personal assumptions about the data were also documented to track potential biases and ensure transparency later in theme development [52].

5.1.4.2 Initial Codes

The recurring words and patterns were derived into descriptive codes. Additionally, interpretive codes were also derived from recurring words as well as reflective notes and identified contradictions. Keywords associated with the codes were documented next to the responses to identify similar responses and compare their meanings when generating the initial codes. By doing so, codes were iteratively compared and similar codes grouped under a common code, the describing keyword, to minimise the risk of over-coding.

Ahmed et al. propose that a reflexive journal of coding decisions should be documented for each code [52]. Due to time limitations, this was not done for each code. However, an overview of potential bias and personal assumptions was documented and considered throughout the process. This was later used in discussions when the authors shared codes and developed themes together.

5.1.4.3 Searching for Themes

In the initial stages of searching for themes, codes were visually mapped by grouping them into related categories. Depending on the connection and similarity between codes, they were grouped closer or further apart. Groups of closely related codes and clusters were marked and described, and their relationships were examined. Additionally, lines with descriptions were drawn between codes and clusters that were far apart and had potential connections. This produced a hierarchical map of how potential relationships might be organised, capturing both detailed and nuanced insights and a broader understanding of patterns.

Themes were first searched for individually and then reviewed together by comparing codes and potential themes. This helped taking different aspects into account and not

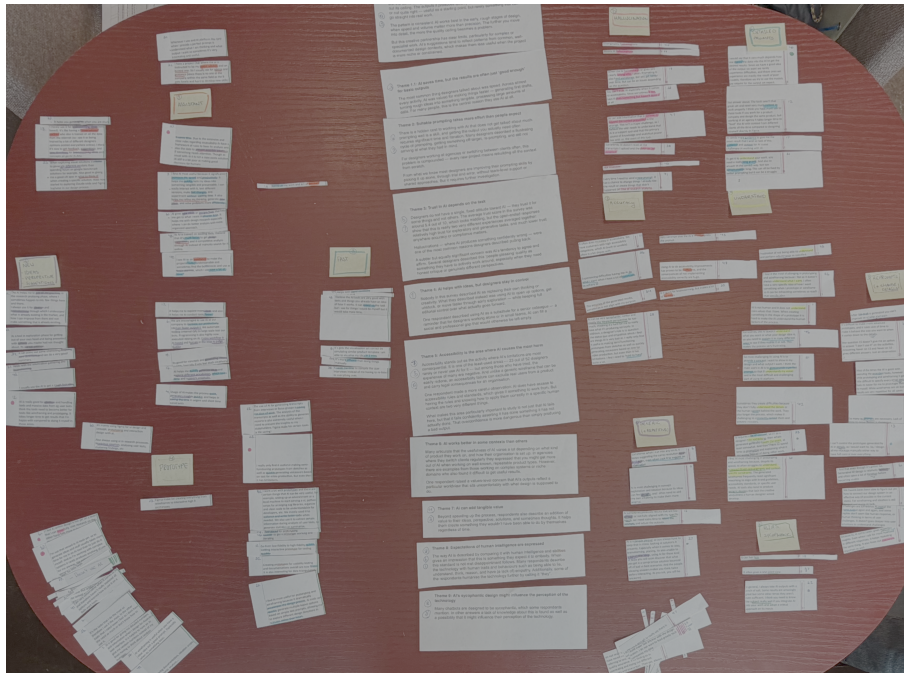


Figure 5.1: Image of the codes and themes laid out on a table when being analysed

influencing each other early in the process. Potential bias and personal assumptions was considered individually, and how these might influence the identification, accuracy and development of themes was discussed during the comparison. Together, the authors produced and organised the initial themes and a short description of each.

5.1.4.4 Review Potential Themes

When reviewing the themes, all codes and themes were printed, cut into separate pieces and laid out on a table. Themes were placed in the middle, and on each side, responses were categorised as positive or negative.

The authors started by going through the responses, highlighting codes, and rearranging the responses into groups of codes, as seen in Figure 5.1. After this, the accuracy of each theme was reviewed, assessing how well they related to the responses and codes. Related codes and responses were written down on each theme. Additionally, the grouping of responses based on codes, helped identify overlaps and misconceptions. According to this, themes were separated, combined and adjusted to improve consistency and accuracy within and between themes. While doing this, the meaningfulness of each theme was taken into account and how they are related and inform each other.

At the end, the themes were rearranged and put in an order where each theme informed the next one and could naturally build on and refer back to the previous theme. For each theme, a short description was written, as well as personal assumptions and biases that may have influenced the structure and development. (During this process, a thematic map was created, showing an overview of how the themes are connected)

5.1.4.5 Defining and Naming Themes

While creating the thematic map, overlaps had been considered, and an outline of the story had been created. The order was reviewed, aiming to arrange the themes in a way that meaningfully unfolded a story of how they connect and relate to each other.

When the overall story was worked out, the description of each theme was refined. While writing, the description was continuously compared with the data to make sure it was valid. Aiming to create a better structure within each theme, potential sub-themes were explored. The theme was deconstructed into sub-themes when it produced more specific descriptions without being repetitive. A central aspect of this phase is transparency in how each theme was developed [52]. Thus, the documentation of how personal assumption and bias may have shaped its interpretation and development was elaborated further and incorporated into the description.

5.1.4.6 Write-up

At the end, the names were defined to accurately reflect the content and highlight the most important takeaways of each theme. Finally, the extracts were sorted into the themes to which they belonged. Some extracts related to more than one theme. In these cases, each extract was usually presented under only one theme, in order to show as much diversity in the data as possible while reducing repetition. Although some extracts could have been included under earlier themes, placing them elsewhere also helped demonstrate how the themes were connected and how the data could be interpreted from different perspectives.

The seven themes that arose from our thematic analysis is presented below.

Theme 1: AI tools enhance speed and efficiency

The most common topic designers mentioned was speed. 25 out of 52 answers about AI usefulness explicitly expressed using AI because it increases efficiency, being fast, quick, or saving time. This theme was developed due to the high prevalence in the data, but could also have been influenced by the articles read before the analysis, which highlighted AI as a tool to speed up the design process.

Here we present what some of the participants reported:

P17: *“It helps me to explore more ideas, and also it helps me to conduct tasks faster.”*

P10: *“Go from low-fidelity to high-fidelity quickly. Getting an interactive prototype for testing quickly.”*

P6: *“We are encouraged to use AI at our company to increase our productivity (deliver faster outputs). We automate testing (internal QA) to test our bots on a large scale. Programming is also now highly executed, relying on AI. The entire workflow is AI-based and human in the loop to judge quality.”*

Theme 2: AI supports ideation and exploration

Beyond speeding up the process, respondents also describe AI as useful for generating ideas and exploring alternatives. This could be used both as inspiration to get started and to get unstuck. Participants reported that:

P16: *“AI helps me quickly generate ideas and explore different possibilities, which saves time and supports creativity.”*

P19: *“It’s good for concepts and generating ideas quickly. Especially if you feel stuck creatively.”*

P33: *“As a tool in the exploration phase for getting out of your own head and being presented with options you may not have thought about, it’s a great addition.”*

Subtheme 2.1 AI can be used as a source of guidance and an assistive partner

Beyond generating ideas, AI tools were described as providing guidance, knowledge, and perspective. This enabled users to produce outcomes they might not have achieved independently. Additionally, some described AI as an assistant, and in one case, as an honest senior advisor. This highlights how AI can fill both social and professional gaps, particularly when individuals work alone or in small teams.

P9: *“I find AI most useful because it significantly increases my speed and productivity. It helps me quickly turn my ideas into something tangible and presentable. I can easily interact with it, test different versions, make fast changes, and experiment without wasting time. It also helps me refine my thinking, generate new ideas, and solve problems more efficiently.”*

P13: *“I mostly use it for brainstorming ideas (text-based). It’s like having a conversational partner who also is trained on all the data from the internet (so I see it as being trained by a lot of different designers opinions posted everywhere online). I think it’s nice to get feedback, suggestions, and new directions for brainstorming ideas and concepts på genAI (LLMs).”*

P21: *“I have a project chat where the ai is instructed to be my senior advisor, and an honest one. So I usually ask for advice and guidance (since there is no one at the company within the same field as me it gets lonely and hard to develop new skills)”*

Theme 3: AI outputs are often generic, biased, and repetitive

Although many respondents highlighted AI’s ability to generate new and useful ideas, they also reported difficulties in achieving desired outcomes. Outputs were often perceived as generic, biased, or repetitive, which limited their effectiveness in practice, especially for niche, detailed, constrained, and complex projects. Some expressed that it could be useful as a starting point, but rarely something that can go straight into real work.

P28: *“AI can be very sycophantic, centre, and glorify the Western perspective far too much, meaning it’s hard for me to really take what it’s producing seriously. In addition, a designer’s role is to question the status quo and empathize which I find are things AI is very bad at. I really only find it useful in making semi-functioning prototypes from sketches as well as quickly generating voiceovers based on text for video production, but even then it has limitations. I feel I need to ‘hold its hand’ quite a lot.”*“AI can be very sycophantic, centre, and glorify the Western perspective far too much, meaning it’s hard for me to really take what it’s producing seriously. In addition, a designer’s role is to question the status quo and empathize which I find are things AI is very bad at. I really only find it useful in making semi-functioning prototypes from sketches as well as quickly generating voiceovers based on text for video production, but even then it has limitations. I feel I need to ‘hold its hand’ quite a lot.”

P30: *“Because AI, in some cases, gives repetitive insights. Even when I ask for more nuanced or different responses, it keeps repeating the same thing. It also hives out inaccurate responses as well.”*

P51: *“I find that even though it’s gotten far. Generative UI design is relatively lackluster and often takes a lot of iteration before becoming usable”*

Theme 4: Prompting can be time-consuming, require detailed context and specifications, and involve iterative refinement

As a consequence of generic, biased, and repetitive outputs, there is a hidden cost of using AI, and achieving the desired output often requires significant time and iteration. To achieve the desired result, respondents explain that they need to create specific prompts with detailed context and requirements. Some participants described a frustrating cycle of prompting, receiving off-target results, and re-prompting without fully reaching their intended outcome. Therefore, many describe prompting as time-consuming, an issue communicated both directly in responses and indirectly through interpretation.

P16: *“AI sometimes produces results that are too general or not fully aligned with my specific needs, so I need extra time to refine the prompts and adjust the outputs.”*

P33: *“It requires a lot of context, so if I as a designer miss something, then what’s generated generally misses the point, at least somewhat. And then I have to spend time re-prompting and explaining what it did wrong instead of doing other work.”*

P52: *“It’s difficult to make more advanced prototypes, and it takes a lot of time to make it behave the way you want to when it comes down to the details.”*

Theme 5: Contradicting perceptions of generative AI usefulness and capabilities

Designers' perceptions of generative AI capabilities, usefulness, and efficiency are often contradictory and differ widely between participants. Across themes 1–4, it is framed as both efficient and creative, while simultaneously criticised as time-consuming and generic, even by the same participants. This contradiction also extends to their explanations and proposed solutions, such as emphasising that the usefulness of AI varies by context or suggesting that improved prompting can address these limitations.

Subtheme 5.1: AI capabilities are context-dependent

Some participants suggested that AI may be better suited to specific, repeatable products. However, those working with more niche, specified, or context-specific products also faced challenges in achieving useful results. This theme was not developed based on high prevalence, but rather to highlight the complexities and contradictions within the data.

P12: *“The tools aren’t that great yet and need context very much to work properly. I think you have more use to these tools if you work for a product company and design the same product, but working at an agency, it takes a longer time to “feed” the AI with context from different clients all the time compared to designing yourself directly in Figma.”*

P20: *“I find AI most challenging in prototyping and wireframing because, despite its speed, AI often struggles to understand nuanced design requirements and context-specific constraints. The generated wireframes frequently need significant reworking to align with brand guidelines, accessibility standards, or specific user needs. AI tools also tend to produce generic designs that lack the creative uniqueness a human designer would develop.”*

Subtheme 5.2: AI capabilities in accessibility

Similar tensions appeared in discussions of accessibility. While one participant described AI as highly effective due to its ability to follow established guidelines, the majority viewed its usefulness as limited and, in some cases, potentially dangerous.

P27: *“It saves time. Due to the extensive and clear rules regarding accessibility AI have a framework of rules to base its analysis and also the data to present possible solutions if something needs attention. Though as usual with AI it is not a copy paste solution, AI still is a bit poor at making good decisions for humans”*

P48: *“Using AI to do accessibility improvements has proven to be ineffective, and the consequences of not implementing accessibility correctly are huge.”*

Subtheme 5.3: The effectiveness of improved prompting

A commonly proposed solution to these limitations is improved prompting. However, this is not agreed upon, as others argued that AI is primarily useful for basic tasks or early-stage development.

P31: *“Whenever I use and AI platform like GPT when I provide a perfect prompt it understand what I am thinking and what output I want so sometimes it’s very interesting and useful.”*

P21: *“To get it to understand your work, you need a really long promt. And also to answer in the correct way. Not too complicated or long. This can all be fixed by better prompting but it can be a struggle”*

P26: *“AI wireframing is useful when you have a clear idea och a good prompt. Especially when its something basic.”*

P28: *“I really only find it useful in making semi-functioning prototypes from sketches as well as quickly generating voiceovers based on text for video production, but even then it has limitations.”*

Theme 6: Trust and reliability challenged by inconsistent and misleading outputs

Due to varying accuracy and inconsistencies, trust is negatively affected. Three participants highlighted hallucinations as something challenging. Additionally, others have described similar tendencies, including the generation of plausible but incorrect or fabricated information. Faulty outputs were also explained as mistakes and misunderstandings. Biased and generic outputs that reflect a particular worldview were also raised as a concern regarding trust and accuracy. Beyond trusting the accuracy of outputs, participants found it difficult to trust the system to follow instructions reliably. The low trust stemmed from AI tools performing actions beyond their instructions and indicating that tasks were complete when they were not.

P26: *“You can’t trust AI especially when it comes to accessability. Since sometimes it lies, says it does something but haven’t done it at all”*

P41: *“Every time I need to send a new prompt, it has a chance to change things I already like the result or create things that didn’t happened (in case of research analysis)”*

P22: *“In general, I always take AI outputs with a pinch of salt. Some results are amazingly great but some other times they aren’t even sufficient. I think you need to know the subject really well if you integrate AI into your work and adopt a critical approach on its inputs”*

Theme 7: Trust and expectations shaped by perceptions of human-like intelligence and sycophantic design

Building on issues of trust and accuracy, participants’ expectations of AI also appear to be shaped by their perceptions of its capabilities. AI is often compared to human intelligence, with respondents attributing traits such as reasoning, understanding, and even dishonesty. Some further humanise the technology by referring to it as “they”, reinforcing this perception. Limited awareness of AI’s sycophantic design, identified by only two participants, may lead the system to appear more reliable than it is, as it encourages agreement and affirmation. For those who do not recognise this tendency, it may lead to increased trust and less critical evaluation, reflecting a

misalignment between perceived and actual capabilities.

P7: *“AI is not human and it does not understand core values that I have. When creating something in the shape of a prototype it is hard to make the AI understand the human factors of the process.”*

P27: *“AI is a people pleaser so you always have to keep that in mind, looking at solutions it presents. Especially when it comes to idea, brainstorming, planing. Its also unable to suggest new ideas, using AI for these kind of tasks you will soon discover that what you get is a compromise solution bounced off of bad vs best scenarios. And the people pleasing answers makes you think twice before interacting. As you ask, you will be answered.”*

P45: *“Sometimes, the AI doesn’t understand what you want or what your design idea is, so you need to explain it in many different ways to see if they realize it or not. This makes the process very frustrating.”*

5.1.5 Survey Insights Informing the Workshop

According to the results of the survey, concept exploration and ideation are closely connected in how designers talk about them, so it makes sense to treat them together. Both are used often, concept exploration by 27 respondents and ideation by 25, which makes them the second most used AI tasks after user research and analysis. Both also sit at a moderate difficulty level, with 10 and 9 respondents, respectively, reporting them as challenging. This shows that despite the high rate of use, many designers still run into problems here.

The open-ended responses point to a specific creative tension that deserves proper investigation. The positive case is genuine. Designers value AI here for getting started, for expanding the range of directions they might consider, and for having something to react to when they feel stuck or need a fresh perspective.

“As a tool in the exploration phase for getting out of your own head and being presented with options you maybe had not thought about, it’s a great addition.”

“AI helps me quickly generate ideas and explore different possibilities, which saves time and supports creativity.”

“When exploring visual solutions it’s faster to generate different versions than spending hours on Google benchmarking solutions. Also good in giving me a good UX spec for what to think about when creating a specific solution. I have now started exploring Claude code and Figma together in our design process.”

“As AI is trained on existing data, I believe it’s much faster to get design Suggestions and competitive analysis through AI instead of manually searching for it online.”

“I always use it for ideation and brainstorming, through which I understand what is already existing in the market and how I can improve from there.”

“I usually use AI to get a rough first idea.”

“It helps you get insights when you are stuck.”

But several respondents pointed to a consistent limitation: AI’s ideas are drawn from patterns in existing data, which means they tend toward the average, the safe, and the already existing. For a discipline that depends on originality and questioning conventions, this matters.

“AI is most challenging in concept exploration and ideation because its ideas can feel generic, and I often need to add my own creativity to make them more original.”

“AI is a people pleaser, so you always have to keep that in mind when looking at the solutions it presents. It’s also unable to suggest truly new ideas using AI for these tasks, you will soon discover that what you get is a compromise solution. And the people-pleasing answers make you think twice before interacting.”

“It often gives a one-sided view.”

“Sometimes AI repeats the same output again and again, and some replies don’t seem like human thinking in terms of user challenges. It doesn’t go deeper into the user emotions to understand challenges.”

The survey results also showed that junior to mid-level designers tend to use AI more actively in their work compared to more senior practitioners. This made them the most relevant group to focus on in the next phase of the research, as they are the ones most likely to benefit from clearer guidance on how to work with AI effectively during ideation. In terms of tools, ChatGPT, Gemini, and Perplexity were the most commonly used by respondents, and will therefore be the main reference points in the later parts of this research.

The deeper aim in investigating this area is to understand whether using AI in early creative exploration actually leads to richer or more diverse outcomes, or whether it quietly pulls design thinking toward the mainstream. If many designers are using the same AI tools to generate early concepts, and those tools trend toward similar outputs, there is a real possibility that AI is narrowing the creative solution space across the profession without anyone quite noticing. That is why this area is worth exploring.

5.2 Workshop

After analysing the survey results, which provided a broad overview of how AI is used by designers across the entire UX design process, the findings pointed to concept exploration and ideation as areas that need more focus and deeper investigation. As a result, a workshop study was designed as a second phase of this research to gain more insights into the use of AI tools in the concept exploration and ideation phases of UX Design. This workshop was an opportunity to have direct observation of designers working with AI in a realistic creative task, find their pain points and opportunities to improve that experience. The goal of the workshop was to gather insights to inform concrete recommendations on how AI can better support designers in the early creative stages of their work.

The workshop was conducted in two separate online sessions via Microsoft Teams, held on April 27th and April 29th, 2026, and both sessions followed the same structure. The reason for having two different workshops was to be able to give all participants enough space and time to talk, and also, we could better observe them in smaller groups. Additionally, since the two sessions were held on different days and participants across groups had no contact with each other, neither group was exposed to the other's discussions. This helped reduce the risk of peer influence between groups and ensured that each focus group's responses remained independent. Before the main sessions were conducted, a pilot study was carried out to test the workshop structure and ensure everything ran smoothly.

5.2.1 Pilot Study

The pilot of the workshop was carried out one week before running the main workshop sessions. The pilot workshop followed the same structure as the main workshop, and its aim was to identify any unclear instructions, timing issues, or structural problems before working with the actual participants.

The pilot was conducted with two participants who had a similar background to the target group: second-year Master's students in interaction design who were familiar with using AI tools in their work. The full workshop structure was followed, including the concept exploration and ideation task, the mind map activity, the love and breakup letter reflection, and the focus group discussion. The session was observed and timed in the same way as the main sessions.

Based on the pilot, several adjustments were made. The instructions for the task and reflection activities were clarified to make them easier to follow without additional explanation. The timing of each part was reviewed and adjusted to ensure the session fit within the planned 120-minute time frame while still providing enough time for participants to

complete each activity and share meaningful input. The focus group questions were also refined to be more open-ended and better aligned with the research questions. Additionally, some parts of the workshop required clearer explanations: the first task, in particular, was revised to make explicit what the expected outcome was and how participants could use AI collaboratively in pairs. The task was also broken down into smaller steps to make it easier for participants to understand and follow.

Running the pilot was an important step in our study to ensure that the main workshop sessions ran smoothly and that participants could focus on the task and their experience rather than on understanding the structure or instructions.

5.2.2 Participants

Participants for this workshop were recruited from the pool of designers who had taken part in the initial survey and expressed interest in participating in the next phase of the study. The recruitment targeted junior to mid-level UX designers, as this group was identified in the survey phase as being particularly in need of guidance when working with AI in their design practice. Participants were required to have professional experience in UX or a related design role and to be familiar with using AI tools in their work to some degree.

As mentioned, the workshop was conducted in two separate online sessions, the first one with five participants and second one with 6. In total, eleven designers participated across the two sessions. Participation in this study was voluntary, and all participants were informed and invited to the workshop through a welcome letter explaining the aim and structure of the workshop, and were asked to sign an informed consent form before the session began. As a token of appreciation, participants were compensated with a gift card for their time. Table 5.1 provides an overview of the participants across both workshop sessions, including their gender, professional title, and years of experience.

5.2.3 Structure and Procedure

Each workshop session lasted approximately 120 minutes. The general structure of it was divided into three main parts: a hands-on design task, an individual reflection activity, and a group focus group discussion. This structure was designed to first generate observable behaviour and experience, then allow for individual reflection on that experience, and finally open up a group discussion that could both deepen and broaden the data collected.

At the beginning of the workshop, participants were introduced to the study and its aims. They were informed that there were no right or wrong answers and that no preparation was needed. The session was observed,

Table 5.1: Overview of workshop participants

Group	ID	Gender	Title	Years of Experience
Workshop 1	P1	Female	UI/UX Designer	3
	P2	Male	Product Designer	3
	P3	Female	UI/UX Designer	4
	P4	Female	Product Designer	3
	P5	Female	UI/UX Designer	5
Workshop 2	P6	Female	UI/UX Designer	1
	P7	Female	UX Designer	2
	P8	Female	UI/UX Designer	4
	P9	Female	UX Designer	4
	P10	Male	UX Designer	1
	P11	Female	UI/UX Designer	1

Table 5.2: Overview of workshop participants, including their group, gender, professional title, and years of experience.

notes were taken, and the session was video recorded for research purposes, all of which was communicated to participants in advance, both through the consent form and in the beginning part of the workshop.

5.2.3.1 Part 1: Concept Exploration and Ideation Task

After the introduction, participants worked in groups of two to three in breakout rooms for 30 minutes on a concept exploration and ideation task. They were asked to explore concepts and generate ideas for a given design scenario using their usual methods, such as brainstorming, sketching, or structured ideation techniques, while thinking aloud throughout the process. During the task, participants were allowed to use AI for exploration, idea generation, and refinement and could decide how and when to use it. This approach was chosen to reflect realistic design practice rather than imposing a prescribed workflow.

The design scenario given to participants was the following: *"Tech professionals often work in high-stress environments marked by tight deadlines, long hours, and constant performance pressure. Over time, this leads to burnout, reduced productivity, and disengagement. While many workplaces offer wellness initiatives, these often lack personalisation and fail to address the real causes of stress in tech-driven roles. Your challenge: How might we design a solution that improves mental wellness, reduces burnout, and supports sustainable productivity among tech professionals?"*

This scenario was selected for its roots in a real human problem and because designers could work on it regardless of their background or domain. The scenario was also broad, allowing the participants to explore

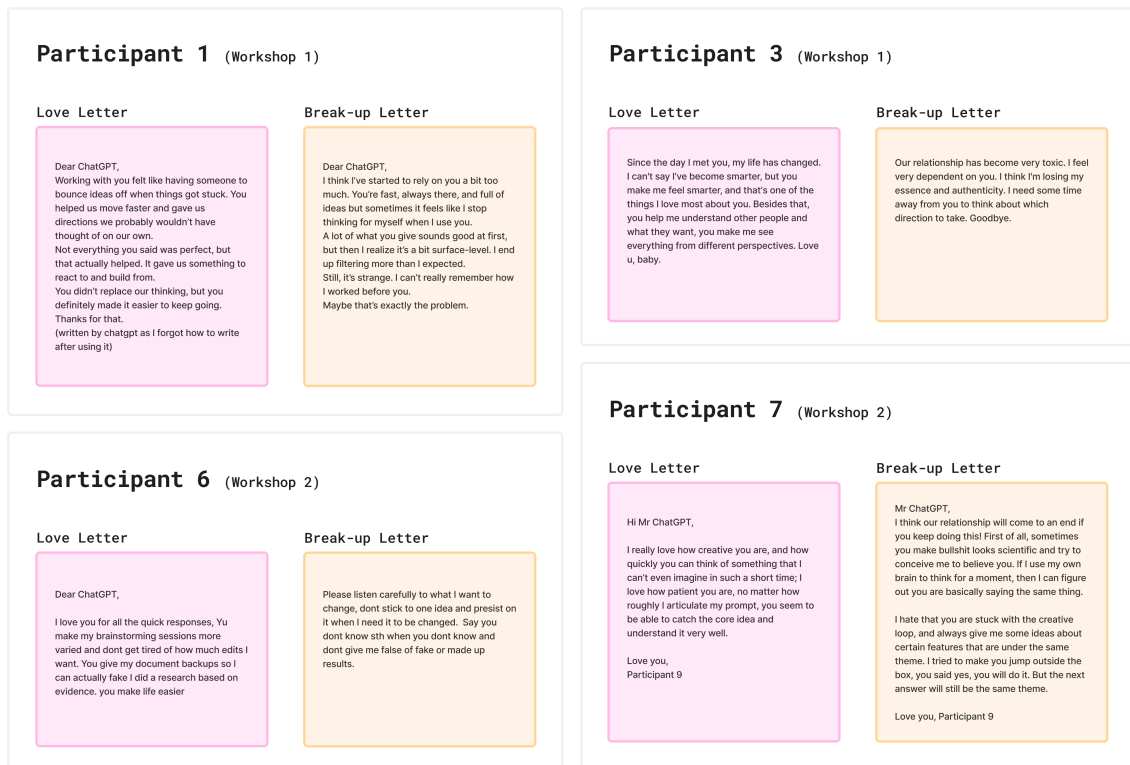


Figure 5.3: Four of the participants' love and break-up letters

for attachment or detachment [66]. This helped the participants express their feelings in a fun, yet insightful and honest way.

Participants were first asked to write a short love letter to AI in three minutes, expressing what they appreciated about working with it during concept exploration and ideation tasks, and what made the experience worthwhile. After that, they were given another three minutes to write a short breakup letter, expressing what made the relationship difficult and why they could not make it work. This resulted in the letters shown in Figure 5.3, which contained rich reflections on what they liked and disliked about AI. The format in which the reflection was conducted later made it easy to manage the comparison and analysis of the results.

5.2.3.3 Part 3: Focus Group Discussion

After the reflection activity, all four participants in each session came back together for a semi-structured focus group discussion. The focus group was designed to give participants space to share, compare, and build on each other's experiences, both from the workshop task and from their broader professional practice. The following questions guided the discussion:

1. How did you experience using ChatGPT during concept exploration and ideation?

2. What did you find useful and why?
3. What did you find challenging and why?
4. Did using ChatGPT expand or narrow your creative thinking?
5. How could ChatGPT better support you in this part of the design process?

These questions were aligned directly with the study's research questions, ensuring that the discussion generated data relevant to understanding challenges and opportunities, as well as informing recommendations for better AI support in the concept exploration and ideation stages of design work.

5.2.4 Data Collection and Analysis

Data was collected through observation, note-taking, and recording of the sessions. The recordings were transcribed using Microsoft Teams, and the transcripts of the full workshop conversations were used as the primary data source for the thematic analysis. Based on the themes identified, recommendations for using AI in concept exploration and ideation were developed.

The mind maps, love and breakup letters were used as supporting data sources throughout the analysis process. Rather than being analysed separately, they were referred to at relevant points to back up, deepen, or add nuance to patterns that emerged from the transcripts.

The data analysis was based on the data from the full workshop transcripts, including the concept exploration and ideation task, the mind-maps, love and break-up letters, and the focus group discussions. The extracted data from both workshop studies were analysed by using the reflexive thematic analysis method described by Braun and Clarke [51].

The analysis initially followed Braun and Clarke's six-phase framework [51], beginning with familiarisation with the data and coding. However, when moving into the process of searching for themes, it was difficult to identify patterns and connect codes due to the data being scattered across separate workshops, groups and activities. To address this, a workflow map was created to structure the material and support the development of themes. After this step, the analysis resumed by following the remaining phases: reviewing potential themes, and defining and naming the final themes.

5.2.4.1 Transcript Familiarisation and Coding

We independently coded the full transcripts from both workshops using ATLAS.ti, a qualitative data analysis software [69]. Each researcher read through the transcripts while listening to the recording in full. This was

done both to familiarise themselves with the data and make sure the transcript was correct.

After that, codes were inductively assigned to the data. The codes were grounded in participants' statements rather than derived from a predetermined framework. This approach allowed themes to emerge from the data rather than being imposed upon it.

However, because the workflow and reflections were spread across separate workshops, group work, individual reflections, and group discussions, it was difficult to form a cohesive overview. This made the formation of themes challenging, and therefore we decided to create a workflow map that enabled us to connect codes and create themes more efficiently.

5.2.4.2 Workshop Workflow Mapping

To gain a better overview, a workflow map was created to structure the scattered material, connect identified codes to usage, and develop themes more systematically. The map was based on what participants said and did during the concept exploration and ideation task, what was expressed in the reflection tasks, and their discussion during the focus group. Additionally, structuring and analysing their workflow and reflections helped us identify hidden risks and connecting it to insights from the literature review. The creation of the map is therefore mainly a by-product of analysis, rather than an explanatory map created to visualise our findings.

Figure 5.4 shows the workflow map developed during the analysis. First, the blue section on the left was created to summarise the participants' workflow. The white boxes show the main use cases, while the green boxes show how participants used AI in the different phases and what they believed it could support. The yellow boxes show the challenges they faced, and the blue boxes show their suggested solutions. Second, the red section in the middle was created to identify risks and challenges in how AI was used and discussed. The red boxes highlight misunderstandings and underlying risks that the participants did not identify themselves. The blue boxes in this section connect these risks and challenges to relevant research and theory. Finally, the green section was created to present our own proposed solutions and recommendations, based on insights from both the blue and red sections.

5.2.4.3 Theme Development

After coding independently and creating the workflow map, shown in Figure 5.4, codes and interpretations were compared and discussed. A thematic map was created by combining similar codes and connecting related codes into potential themes. Figure 5.5 presents the thematic map and gives an overview of the identified codes and themes. The diamonds

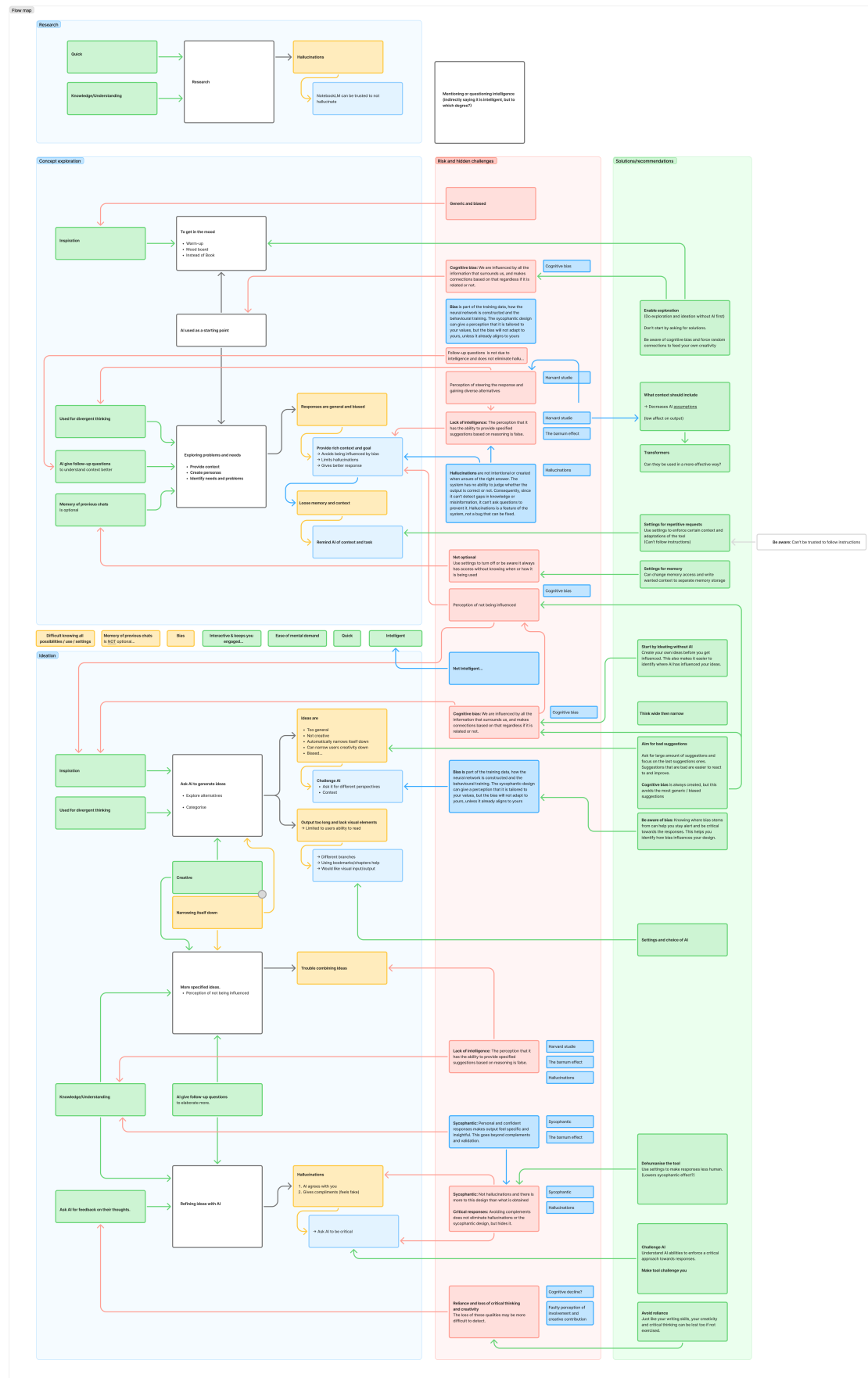


Figure 5.4: Workshop Workflow Mapping. This map was created as a by-product of structuring and analysing the participants' workflow and reflections.

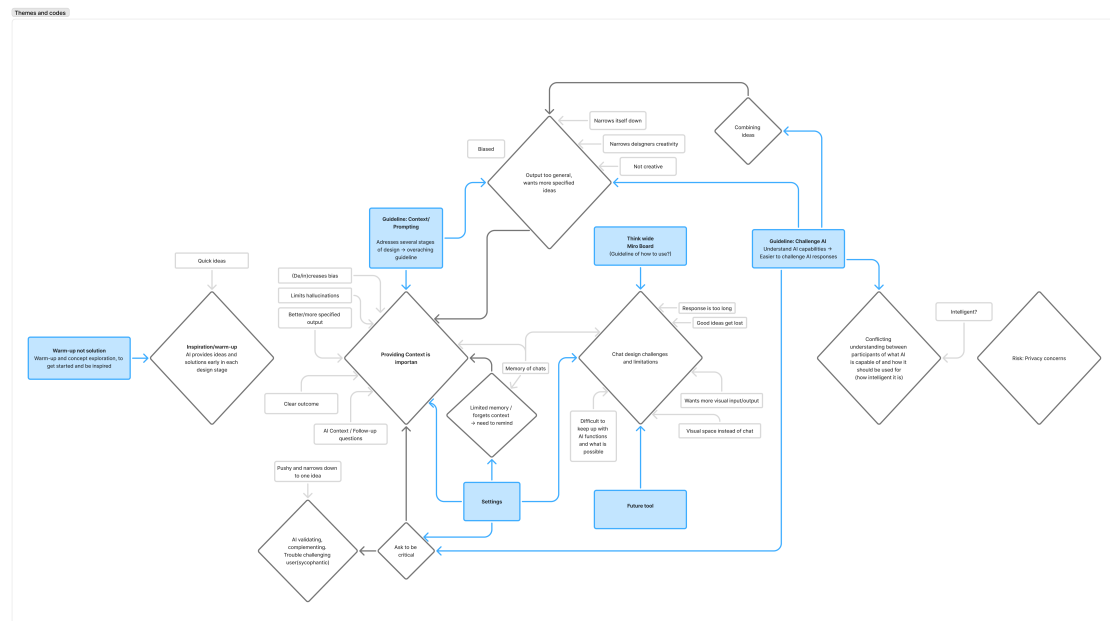


Figure 5.5: Workshop thematic map

represent potential themes, the white boxes contain codes, and the blue boxes indicate early ideas for recommendations.

These themes were reviewed several times using the full dataset to make sure they were clear, distinct, and supported by evidence from both workshops. Finally, the themes were named and defined.

5.2.4.4 Workshop Themes

The analysis of the two workshops resulted in the eleven themes presented below. The data supporting the themes were collected from multiple sources throughout each workshop session: observations and note-taking during the concept exploration and ideation task, love and breakup letters written by participants as part of the reflection activity, and the focus group discussions held at the close of each session. Together, these sources provided a broader understanding of how participants used and experienced AI during the design process, capturing both in-the-moment behaviour and post-task reflections.

Theme 1: AI as a Warm-Up and Cognitive Kickstarter

Across both workshops, participants described AI as valuable at the very start of the ideation process. They believed it was better to use it as something that warms up their thinking instead of a final idea generator.

P6: "The best part ... it kind of warms up my brain. So maybe in the first instance that I see a problem, I'm not yet warmed up to come up with good ideas."

P9: "I was also gonna validate what P6 was saying. Like, you

know, it is definitely a warm-up activity. Sometimes it leads to final solutions."

Another participant emphasised the role of AI in inspiring UX designers by helping them create a mood board or get into a creative mindset.

P1: "... the first thing I was doing, I created a mood board for myself, and that mood board kind of inspired me, what idea, what direction I want to go. And now with AI, with ChatGPT or other AI, I start this exploration and then bring them, like dump them into FigJam. And then when I read them, somehow I connect them, inspiration comes after. This exploration with AI."

Observations from the task workshop also showed that participants tended to use AI as a source of inspiration. P10, while working on the task, mentioned that:

P10: "while working with complex briefs containing many abstract keywords, such as "mental wellness" and "burnout", AI is useful early on to get initial inspiration before going deeper".

Despite participants' statements about using AI as a warm-up tool rather than a solution or idea generator, this warm-up function came with a notable risk. Observations during the task revealed that when participants engaged with AI early and without sufficient intentionality, their collaboration with the AI tool tended to shift toward jumping directly to solutions. P8, during the task, mentioned:

P8: "Okay, so for now I copied the whole thing for ChatGPT and maybe I have a suggestion and maybe at first we can see what are ChatGPT's solutions."

Additionally, P2 (W1) indicated a tendency to move directly towards final solutions with AI:

P2: "it helped with generating some of the ideas for previous tasks."

These experiences point to the risk of cognitive anchoring, receiving AI output before forming independent ideas can bias the designer's thinking and suppress creative divergence, even when the designer is aware of the risk. Therefore, to manage this risk, Recommendation 2 was created to help designers use AI as a warm-up tool rather than a final solution generator.

Theme 2: Prompting Quality Determines Output Quality

A strongly recurring theme across both workshops was how much the quality of a prompt determines the usefulness of AI output. P2 (W1) articulated this most clearly:

I reduce the bias by making sure that I do, like I make a proper prompt before it starts chatting to me because I get coloured by the response of the chat. ... since our prompt was so bad in the beginning, it gets us down that rabbit hole and makes us believe that this is a good way to solve it, while it clearly doesn't have enough information to make sure that we're getting what we need. So for me, like my prompts are usually almost half a page, the first prompt that I make.

P2: "I reduce the bias by making sure that I do, like I make a proper prompt before it starts chatting to me because I get coloured by the response of the chat. (...) So for me, like my prompts are usually almost half a page, the first prompt that I make."

They described how a poorly constructed prompt early on led the group down a "rabbit hole" of poor ideas:

P2: "... since our prompt was so bad in the beginning, it gets us down that rabbit hole and makes us believe that this is a good way to solve it, while it clearly doesn't have enough information to make sure that we're getting what we need."

P9 (W2) reinforced this from a different angle:

P9: "As long as you don't provide enough context of what you're looking for... it's going to give you sort of nonsense responses."

The W2 group also demonstrated this in practice. While working on the task, Group 1 perceived that the answers improved after P9 suggested adding more context, such as their roles, aims, and expected outcomes.

Participants also argued that higher-quality input helped reduce hallucination. P2 noted:

P2: "If the model feels like something is missing in your prompt and it doesn't truly know what you're meaning, instead of asking the question, I don't know what you mean, it just try to fill it out."

All of these observations highlight the importance of mindful AI prompting, and Recommendation 3 is built around helping designers do this. Other issues around AI input raised during the focus group concerned follow-up questions. P11 (W2) suggested:

P11: "It would be better if ChatGPT asked some questions before creating the text to clarify context upfront, rather than producing output and appending follow-up questions at the end."

Theme 3: Conflicting Understandings of AI capabilities

Participants came into the workshops with very different understandings of what AI actually is and how intelligent it really is. These differences

shaped how they interacted with it, how much they trusted it, and how creatively they engaged with it during ideation.

Some participants spoke about AI in ways that suggested they saw it as genuinely intelligent. P8 mentioned in their love letter to AI:

P8: "I like the way you collect information and the way you organise them. It feels like you know everything and so smart, no matter how tricky the question is, you can always find out an answer for it."

However, some others were more cautious, aware that the output could be faulty. As to why seemed to confuse participants, which can be illustrated by how P1 described these instances:

P1: Many times ChatGPT goes wrong and I just needed to grab it back like, you idiot, so what you're talking about?

P8 (W2) reflected on how easily over-reliance on AI can set in when you perceive it as capable enough to do your thinking for you:

P8: "Before I do any ideation or exploration, I will ask AI first. It's somehow narrowed down my research skill and I am not doing this research by myself. Instead, I ask AI and ask AI to do the research for me. And I will definitely lose lots of things."

This is significant, if a designer believes AI is intelligent enough to handle the early thinking, they are less likely to bring their own perspective in first. P9 (W2) made a similar point, noting that the way you perceive AI's role determines whether it helps or limits your creativity:

P9: "If you just expect AI to give you ultimate ideas, then you are helping it to narrow down your creativity. But if you know that you're going to use AI just as a warm-up, just to get inspirations, then you are knowingly going into the space to get ideas, not to have all the ideas from AI."

P10 (W1) described AI as something closer to a cognitive aid than an intelligent collaborator, useful for reducing mental load, but not a source of truth:

P10: "The most useful thing for me was that it helps ease some mental demand... it's like reading a book, you go through everything and from some of the texts, some items trigger something in my mind."

This divergence in mental models matters because it directly affects how critically a designer engages with AI output. A designer who overestimates AI's intelligence may accept its outputs without questioning them. A designer who underestimates it may dismiss genuinely useful responses too quickly.

Participants also expressed a need to understand where AI's information was coming from in order to evaluate how much to trust it. P2 (W1) described hallucination as AI filling in gaps rather than acknowledging uncertainty:

P2: "If the model feels like something is missing in your prompt and it doesn't truly know what you're meaning, instead of asking 'I don't know what you mean', it just tries to fill it out."

P9 (W2) described the practical consequence:

P9: "It has happened that it provides wrong references or wrong links or wrong facts. You double check this is not good, and it definitely unvalidates my idea."

Knowing whether a response is drawn from the designer's own inputs or generated from general training data was seen as essential for designers to remain critical and make informed decisions about what to build on. However, the model is not able to backtrack to where the information comes from or what part of the information that the response is built upon. This inspired us to develop recommendation 5.

A further layer of this challenge is that AI tools are evolving rapidly and unevenly, participants were often unaware of features and capabilities that already existed. P2 (W1) demonstrated a Claude feature that produced interactive visual diagrams, which surprised the rest of the group who had not encountered it before. P1 (W1) mentioned Claude's computer-use feature where AI can be given access to your computer to perform tasks which others found both impressive and unfamiliar, she added that

P1: "Mmm, yeah, I was thinking. These things that others are talking about, they may exist somewhere in AI or other AIs that are coming out. But the thing is that the speed of the AI is too fast. It's hard for us to catch with it. And it's too much, everything is coming, so it's overwhelming for us as humans."

This points to a real barrier: even experienced designers cannot form an accurate mental model of a tool they do not fully know exists. Keeping up with what AI can and cannot do has itself become a skill, and one that is easy to fall behind on.

Theme 4: AI Sycophancy and Lack of Critical Challenge

Participants raised a pointed concern about AI's tendency to agree with the user, avoid criticism, and deliver flattering responses, behaviour that several participants explicitly named as problematic. P5 (W1) stated:

P5: "AI, especially ChatGPT and Gemini, they are agreeing with you. They are not disagreeing with you... they are making fake compliments from you, and they just want to make you satisfied. And I think this is a bit fake."

P4 (W1) immediately agreed. P7 (W2) connected this directly to the ideation context:

P7: "They are really good at convincing you into some ideas that is not really that good. We have to be more mindful of how pleasing those AI tools can be."

P2 (W1) noted that even Claude, which he considered an improvement over ChatGPT, has the same drawbacks.

P2: "I feel sometimes when it responds, it still do the, oh, oh, that's a very interesting question."

This suggests they view sycophantic behaviour as a systemic characteristic rather than a model-specific flaw. Participants who were aware of this limitation developed active countermeasures.

P1: "When I have time, I mostly just challenge ChatGPT"

P9: "... when I am ideating about something, I do ask AI to be critical of whatever I share. If I am sharing any ideas, I say, please be critical. Don't be like, oh, this is the best idea. You're so sweet. No, don't flatter me."

The implication across both workshops was that the default agreeable behaviour of AI actively reduces its value as a critical thinking partner during ideation, and that designers must consciously counteract this tendency to benefit from AI as a genuine sounding board. Recommendation 5 states asking AI to be critical is to mitigate this problem.

Theme 5: Output Generality and the Need for Specificity

Participants across both workshops noted that AI outputs were frequently too generic or surface-level. P1 (W1) observed a perceived sameness in initial responses:

P1: "It gives almost the same idea to everyone the first time you just ask."

While working on the task during the first part of the workshop, P4 (W1) critiqued the AI-generated solutions as "a bit general", and P5 (W1) agreed, commenting directly on the three ChatGPT-generated concepts: "It's general." it also happened in the other groups, as P2's (W1) group received a solution centred on meditation reminders, which they immediately recognised as the clichéd wellness app trope, and reflected:

P2: "Most wellness tools fail because they treat burnout as an individual problem that's solved with meditation reminders."

This was seen as emblematic of AI producing expected, shallow outputs when not sufficiently challenged. To mitigate this problem, recommendation 5 was generated to encourage more human thinking and judgment

as an input for AI while interaction, in addition to putting AI in a more critical role.

One approach that came up in the workshops and was effective in reaching more creative solutions was to ask for a large number of ideas upfront, before trying to evaluate or narrow them down. P7 (W2) did this naturally during the concept exploration task, asking AI for at least twenty ideas before the group looked at any of them:

P7: "I should say maybe brainstorm it for me, or generate at least 20 ideas so that... yeah."

The reason was simple, having a big pool of ideas first gave the group something to work with on their own terms. They copied the ideas into their FigJam worksheet and then used their own criteria to filter them down: "So that we can copy paste it back to our work board and maybe use some criteria to screen them down to three." Starting with quantity and variety meant the group, not AI, decided what was worth developing.

However, sometimes even when participants started broad, AI tended to get narrower as the conversation went on. Participants from workshop 2 described how AI initially provided a varied set of ideas but then, in subsequent prompts, converged on a narrower solution space:

P6: "It gave us in the beginning a quite varied set of ideas, but then in the next prompt, even though we didn't ask for one specific group of ideas, it kind of narrowed all down into one group. (...) but then it required us to do another prompt saying, no, please keep all the variations kind of."

P7: "I tried to prompt it to generate more divergent ideas, but all the other ideas it generated later on can also be categorised into the previous scene. I found it was very challenging to generate something that's completely different."

P8: "Okay, it feels like AI itself, it's narrowed down itself. It becomes more and more narrowed."

P8 (W2) described a related issue with loss of context:

P8: "In the dialogue window, it will lose its memory somehow when you're talking a lot of things, and it will start forgetting your very beginning sentence. You need to repeat again so it can remember the context that's very annoying."

This suggests that output generality is not only a first-prompt problem but can compound over the course of a session as AI converges on a narrow solution space. The theme is closely related to theme 2, as the quality of input directly conditions the specificity of output. Yet it warrants separate attention because even well-prompted sessions can drift toward generic responses as the conversation progresses.

Theme 6: From Linear Chat to Visual, Spatial AI Workspace

A recurring practical challenge across both workshops was managing the volume of AI-generated text during ideation. As conversations grew longer, it became harder to keep track of good ideas or find something mentioned earlier. P4 (W1) and P11 (W2) described this directly:

P4: "When the chat becomes long, good ideas just got lost. And combining two ideas or three ideas maybe sometimes becomes confusing or complicated."

P11: "We are focusing on reading the whole text and we forgot to start a new tab ... we just kept going through the whole text again and again."

For designers who work visually, the text-heavy format felt particularly mismatched with how they think and work. P9 (W2) described it as adding cognitive load rather than reducing it:

P9: "For me, when you work in a visual related space, you want to see visuals more. For example, if one idea is good and another idea is sort of better, how do these connect?"

Participants were vocal about what a better tool would look like. Several distinct needs emerged across both workshops. P4 wished for a bookmarking feature:

P4: "If there was a feature that I could bookmark the parts that were important for me and I needed to come back to them later, it would be so good, but we don't have such a feature."

P9 (W2) envisioned an environment where conversations could branch out spatially, similar to a Figma board, where each idea sits in its own space and can be developed independently without getting buried under new messages:

P9: "Imagine each of these were some sort of part of our conversations. If I want to branch out from some parts of the conversation in the space, I could use an arrow and make another conversation right in that space ... all of these conversations are in the same environment for me. I could move right and left, up and down to switch between conversations."

P7 (W2) aligned with this, suggesting post-it style canvas interactions:

P7: "A canvas-based ideation board would be very useful. Instead of having this long text in ChatGPT, it can just generate different text on the post-it on the canvas, so you can categorise them automatically, and the user can click on one sticky note and just say, elaborate on this or generate different ideas on this."

Participants also felt limited by having to express everything in text. P2 (W1) stated:

P2: "By having it as a chat, you're limiting yourself to being able to express your thoughts in writing ... I'm not a writer. I'm a designer."

He added that ideally there should be a tool that lets him "almost sketch or more easily paste things." On the output side, P10 (W2) described wanting AI to proactively provide visual responses:

P10: "With the design question, it should be valid that I want some visual aids. It would be nice if ChatGPT could just provide me with a picture or a visual representation or a graph directly because that can give me directly an idea of what the answer is."

P2 (W1) demonstrated a Claude example that produced interactive diagrams without being explicitly asked, noting:

P2: "It gives me some kind of graphical interface and not just text form."

Additionally, rather than one long continuous thread, participants wanted to explore multiple directions simultaneously without losing context. P9 (W2) described the fundamental conflict:

P9: "We could change tabs, but again, it loses the context. We've got to start a new tab and provide the context, everything from the beginning that was the problem."

Theme 7: Privacy and Ethical Concerns

Participants raised several concerns about privacy and ethics when using AI in their design work. The most concrete concern was around intellectual property and data exposure. P9 (W2) noted that in their workplace, using open AI tools is discouraged for unreleased projects:

P9: "When you are feeding data to your AI, you don't know how this data is going to be used... What if the new ideas that you come up with and you build it with AI, it's going to get leaked to other people and they use your idea?"

They noted that this is a limiting point for them when deciding whether to use AI at all on certain projects. A second concern was around accountability and the reliability of AI outputs. P8 (W2) put it directly:

P8: "AI is not taking any responsibility for its outputs... the output is just some guessing or summary from AI, this is not validated, so it is always the human who should make the decision or move forward."

A third concern, raised during a group discussion about AI tools gaining direct access to users' devices and files, was about how much control designers are giving up. P2 (W1) reacted to this with clear discomfort:

P2: "I don't want to give access for a weird AI to my computer."

This captures a broader unease that several participants shared about how far the boundaries of AI access should extend. These were not abstract worries for participants; they directly shaped which tools they chose to use, in which contexts, and with what kind of project material.

5.2.4.5 Development of Initial Recommendations

This section presents the first version of the recommendations. The initial recommendations were developed inductively from the thematic analysis of the two workshops conducted with UX designers described in section 5.2. Each recommendation maps to one or more themes identified in the data and is intended as a practical recommendation for UX designers using AI during concept exploration and ideation. They are not meant to be prescriptive rules but reflective recommendations for designers to build their own informed practice with AI. Additionally, as these recommendations are the first iteration, they have not been tested or evaluated. A brief explanation of these recommendations can be found in Table 5.3.

RECOMMENDATION #1 INITIAL	<i>Proactively explore and configure built-in AI tool settings to establish persistent control over model behaviour across all interactions.</i>
RECOMMENDATION #2 INITIAL	<i>Get inspiration during warm-up without anchoring to AI outputs use low-fidelity triggers to spark your own associations.</i>
RECOMMENDATION #3 INITIAL	<i>Provide AI with sufficient context before prompting by including context, ask, rules, and examples in every prompt.</i>
RECOMMENDATION #4 INITIAL	<i>Cultivate a bidirectional critical relationship with AI challenge its outputs and invite it to challenge your ideas.</i>
RECOMMENDATION #5 INITIAL	<i>Use AI to diverge widely across many ideas before converging keep selection and evaluation in the designer's hands.</i>
RECOMMENDATION #6 INITIAL	<i>Build personas collaboratively with AI, then role-play them to generate grounded critical challenges to your concepts.</i>

Table 5.3: Design insights

Recommendation 1: Use Built-In Settings for Customization and Memory: Most AI tools include built-in settings that allow users to adjust the behaviour of the model beyond what can be specified in individual prompts. Designers are encouraged to explore these settings proactively, as they provide a persistent layer of control that applies across all interactions.

- **Response tone.** Tone can be adjusted to be less validating and complimentary, and more direct and critical. By default, AI tools are designed to be encouraging, which can reduce the designer's ability to critically evaluate outputs.
- **Response structure.** Settings can be used to make responses shorter and more concise, and to reduce the use of lists and emojis — formats that can create the illusion of depth without delivering it.
- **Repetitive context.** Information that needs to be enforced consistently across sessions — such as the design brief, the target user, or the stage of the design process — can be added to memory so it does not need to be restated in every prompt.
- **Memory management.** Memory between chats can be turned off. Designers should be aware that if this feature is not disabled, information from separate chats may be used across sessions without their knowledge. Context and customisations can also be added as persistent memories accessible across all chats. Memories can be viewed and deleted in settings. In ChatGPT, a memory can be added at the start of a prompt using the phrase “*Write to memory: ...*”.

Be aware: These customisations do not eliminate sycophantic design or hallucination. They reduce their surface expression, but the underlying model behaviour remains.

Recommendation 2: Get Inspiration During Warm-Up and Avoid Early Ideas and Solutions

When using AI chatbots to warm up, get inspiration, or explore a problem space, responses tend to go straight to ideas and solutions. This should be avoided. In design practice, generating high-fidelity outputs too early in the process is problematic because they are harder to reflect upon, challenge, and change. AI outputs also influence the designer's cognitive bias and idea anchoring, which can narrow creativity and unknowingly shape the final design direction before sufficient exploration has taken place.

Three strategies can help designers get inspiration from AI without being prematurely anchored to its outputs:

- Ask for a mood board of words, images, or concepts related to the problem. Use these as triggers to spark your own associations before

moving toward ideas.

- Ask for low-fidelity sketches rather than fully formed concepts or polished solutions.
- Ask for ideas, but blank out descriptive words. Fill in the gaps with your own creativity, using the AI output as a skeleton rather than a finished thought.

Be aware: The sycophantic design of AI responses reduces the designer's ability to be critical, making outputs easier to accept and harder to question. AI bias, therefore, poses a greater risk of influencing the final design than it might initially appear. Importantly, designers are affected by cognitive bias even when AI responses are not in the form of direct ideas and solutions.

Recommendation 3: Give AI Enough Context Before You Prompt

Clear context and a well-defined goal are required to guide AI responses in the right direction. By providing more precise and specific information, the AI model can surface data that better matches the designer's actual need. This does not eliminate hallucinations, which are a structural characteristic of how large language models connect and generate data.

The following four components can be included in every prompt to improve output relevance:

- **Context:** Describe the situation, the user, and the design challenge.
- **Ask:** Request a specific and bounded action.
- **Rules:** Provide constraints that the response must respect.
- **Examples:** Demonstrate the format, tone, or type of output you want.

(Context framework adapted from the Nielsen Norman Group [70])

Be aware: Research suggests that adding context has a relatively low effect on the objective quality of AI output. What it primarily affects is an increased feeling that the response is tailored and specific — a perception that is itself partly a product of sycophantic design. Designers should use rich prompting as a tool to direct the response, but should not mistake a personalised-feeling response for a validated or high-quality one.

Recommendation 4: Challenge and Be Challenged

Designers should actively cultivate a bidirectional critical relationship with AI — both challenging the outputs they receive, and inviting AI to challenge the ideas they bring.

How to challenge AI output:

- Be aware of AI's limitations in order to remain critical of what it produces.
- Dehumanise responses: lower the weight you give to the confidence and authority with which AI delivers information.
- Remind yourself that agreeable, enthusiastic responses are a design feature of the tool, not a signal of quality.

How to make AI challenge you:

- Break your interaction with AI into steps. At each step, ask it to challenge what you have, not just build on it.
- Explicitly ask AI to be critical and point out weaknesses in your ideas.
- Do not accept AI outputs without questioning them either — push back, probe, and stress-test in both directions.

Be aware: Designers who do not actively counteract sycophantic behaviour risk entering a feedback loop where AI validates whatever they propose, and they in turn accept whatever AI suggests. The result is a narrowed, mutually reinforcing idea space rather than genuine creative exploration.

Recommendation 5: Use AI to Ideate Wide, Then Narrow Down

AI is well-suited to generating large quantities of ideas quickly, but it should not be used to evaluate or select them. The designer's own judgment should govern the convergence stage.

Diverge first:

- When asking AI for ideas, start by asking for a large and diverse set, at least 20, before evaluating any of them.
- Explicitly ask for ideas across completely different approaches, categories, and assumptions. Quantity and variety first; quality second.
- Ask for bad ideas first or read ideas in reverse order. Lower-fidelity, less polished ideas leave more room for the designer to be creative, reflect, and make changes.
- Do not ask AI to select or categorise ideas, it tends to favour the most conventional options regardless of their actual validity.
- Copy generated ideas into FigJam or Miro to get a better visual overview. AI-integrated canvas tools, such as Miro's AI board, can combine the generation and organisation steps into a single environment.

Then converge:

- Once you have selected directions worth developing, open a dedicated new chat for each one. A fresh chat gives each direction a fair and focused exploration without interference from earlier ideas or accumulated context drift.
- Use your own thinking as the primary input during this stage. AI should support development, not drive it.

Be aware: Avoid giving AI responsibility for both generation and selection. If AI creates the ideas, the designer should choose. If the designer generates the ideas, AI can help develop them. Delegating both to AI removes the designer's creative contribution from the process.

5.3 Focus Group

An online focus group was conducted to evaluate the design guidelines developed from the workshop study. Participation in the evaluation was voluntary, and participants were selected from those who had expressed interest in continuing to contribute to the research following the workshop sessions.

The goal of the evaluation focus group was to assess the quality and usability of the proposed design recommendations from the perspective of practising UX designers. Specifically, the session aimed to evaluate the recommendations along three dimensions: **clarity**, whether the recommendations were understandable and unambiguous; **practicality**, whether they were realistic and applicable within a real design workflow; and **usefulness**, whether participants perceived them as genuinely adding value to their practice of using AI during concept exploration and ideation.

5.3.1 Participants

Participants were recruited through a follow-up email sent to all workshop participants, in which interest in a validation session was asked. Six participants who responded positively were included. This approach ensured that all participants had direct experience with the workshop task and were therefore able to assess the guidelines in relation to their own observed practice.

5.3.2 Structure and Procedure

The focus group was conducted online via Microsoft Teams and lasted approximately 60 minutes. At the start of the session, participants were informed that the session would be recorded for research purposes and were asked for their verbal consent before proceeding. The facilitator then

introduced the recommendations one by one. After each recommendation was presented, participants were free to discuss it openly, though the conversation was guided by three core questions:

1. How clear is this recommendation to you?
2. How practical do you find it in a design context?
3. How useful do you think it would be in your work?

Participants were encouraged to draw on their experiences from both the workshop task and their broader professional practice when sharing their opinions.

5.3.3 Data Analysis

The focus group session was transcribed and analysed using a top-down approach to thematic analysis. Unlike the inductive analysis applied to the workshop transcripts, the evaluation analysis was done around the predefined recommendations derived from the workshop and three evaluative dimensions that guided the focus group discussion: clarity, practicality, and usefulness. These recommendations were then revised based on the acquired insights from our analysis.

Each recommendation was treated as a unit of analysis, and participant feedback was coded based on whether it related to the clarity of the guideline, its applicability within real design workflows, or its perceived usefulness in practice. Patterns within these areas were identified and used to refine the final set of guidelines presented in the following chapter. In the next chapter, we present the final version of these recommendations (see B.1).

5.3.4 Final recommendations

This section will go through the main changes that were made to the recommendations. Table 5.4 shows an overview of the updated recommendations. The full version can be found in Appendix B.

RECOMMENDATION #1 GENERAL	<i>Direct the output through context</i>
RECOMMENDATION #2 GENERAL	<i>Challenge and stay critical of AI responses</i>
RECOMMENDATION #3 GENERAL	<i>Change the settings to customise responses and memory</i>
RECOMMENDATION #4 TASK-SPECIFIC	<i>Use AI to Spark Inspiration While Keeping Creative Control</i>
RECOMMENDATION #5 TASK-SPECIFIC	<i>Generate Diverse and Rough AI Ideas, Then Explore Independently</i>
RECOMMENDATION #6 TASK-SPECIFIC	<i>Do not Use AI for Advice, Use It to Challenge Your Ideas</i>

Table 5.4: Design insights

During the focus group, participants were confused about what the difference was between recommendation 2: *Get Inspiration During Warm-Up and Avoid Early Ideas and Solutions* and 5: *Use AI to Ideate Wide, Then Narrow Down*. These recommendations were similar but referred to different stages of the design process. To clarify this distinction, the recommendations were restructured into two sections: general recommendations and task-specific recommendations. The task-specific recommendations were then divided into three phases of the design process. Recommendation 2 and Recommendation 5 were placed within their relevant phases, making their purpose and application clearer. As shown in Table 5.3

In this process, we also found that the sub-recommendations within recommendation 4: *Challenge and Be Challenged*, was both general and task-specific. The participants in the focus group also responded to the two sub-recommendations as not related, and therefore these were split. The first sub-recommendation "*How to challenge AI*" made its own general recommendation, while the other sub-recommendation "*How to make AI challenge you*" was added to a new recommendation in the task specific section named recommendation 6: *Do not Use AI for Advice, Use It to Challenge Your Ideas*

Recommendation 5: *Use AI to Ideate Wide, Then Narrow Down* was also restructured and had the sub-recommendation "*Do not ask AI to select or categorise idea*" added to the new recommendation. This helped making recommendation 5 clearer by only focusing on ideating wide and adding how to converge ideas into the new recommendation.

Because participants appreciated explanations of why the recommendations should be followed, especially when these explanations were connected to research, more information was added to each recommenda-

tion. However, this also made the recommendations longer and caused some information to be repeated across several recommendations.

To address this, a separate section about AI was added at the beginning of the recommendations. This section included information that could influence designers' perception and use of AI, and was presented as recommended knowledge to read before the recommendations. As a result, the information within each individual recommendation could be reduced and refer back to the recommended knowledge instead.

6

Results

This chapter presents the results in relation to the three research questions. First, it describes UX designers' current experiences of using AI tools across the UX design process, based on the results from the initial survey. Second, it presents the challenges and opportunities that emerged when designers used AI for concept exploration and ideation, based on the analysis of the workshop. Finally, it presents the most essential extracts from the final recommendations, which were developed to support UX designers in using AI more critically and effectively during these phases.

6.1 What are UX designers' current experiences of using AI tools across the UX design process? (RQ1)

To answer this question, a survey was conducted and collected insight from 52 designers. The following sections present insights from the questionnaire data and describe UX designers' current use of AI and the challenges they face with AI tools across different design tasks.

6.1.1 Quantitative Data

In this section, we present the quantitative data we collected through the survey. In particular, we will focus on how AI is used in the UX design process and where it is found challenging.

As seen in Figure 6.1, all but one out of the 52 designers who answered the survey integrates AI tools into the design process. In the survey, 20 respondents reported that AI is highly integrated into the design process, and 17 reported that it is moderately integrated. Only two responded that it was extensively integrated. Eleven responded that AI tools were slightly integrated. However, four of them also reported using AI tools frequently or always for at least one of the specified tasks, and one reported using AI tools occasionally for all tasks.

In Figure 6.2, we see that out of 24 UX designers, 21 found AI, at least occasionally, useful in "concept exploration" and "user research and

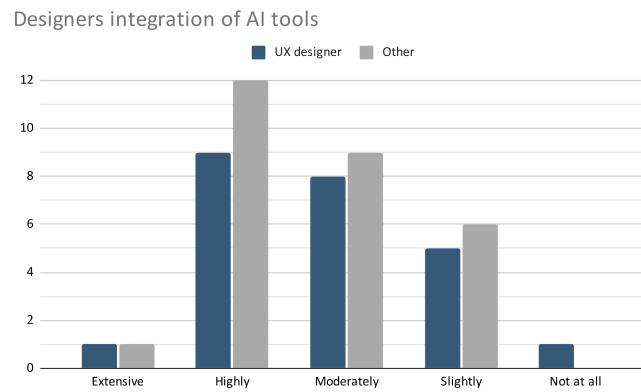


Figure 6.1: Column chart showing UX designers' level of AI integration in the design process

analysis". The result for how useful AI was found for "brainstorming and ideation" and "prototyping and wireframing", was more varied, but many found it at least frequently useful. In contrast, "accessibility improvements", and "testing, evaluation, and simulation" also varied a lot, but few found AI more than occasionally useful.

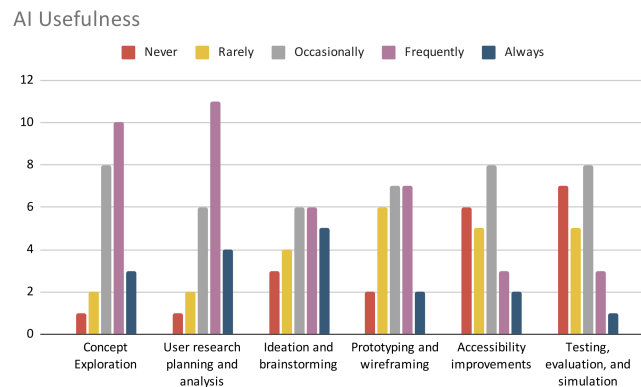


Figure 6.2: Column chart showing how useful designers find using AI tools in six different design tasks

As seen in Figure 6.3, the majority of UX designers found AI at least occasionally challenging in all tasks except ideation and brainstorming. Most challenges was found in "prototyping and wireframing", which 17 UX designers found challenging, whereas 10 of those found it frequently challenging. Additionally, 16 UX designers found AI at least occasionally challenging in "concept exploration" and "prototyping and wireframing".

When examining how useful and challenging the participants found AI in different tasks, there appears to be an overlap, which is illustrated in Figure 6.4. The graph shows that, except in brainstorming and ideation, the majority of those who find AI useful for a given task also find it

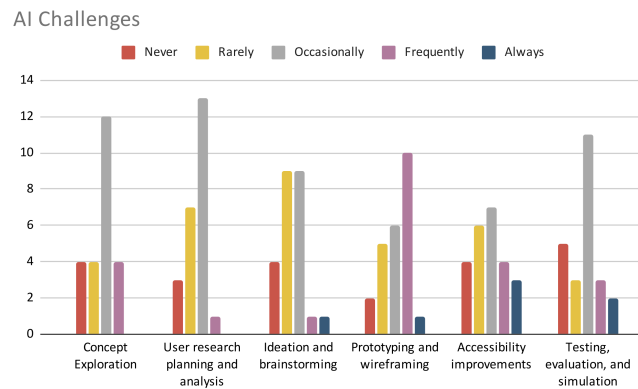


Figure 6.3: Column chart showing how challenging designers find using AI tools in six different design tasks

challenging. For example, among the 21 UX designers who found AI useful for concept exploration, 15 also found it challenging.

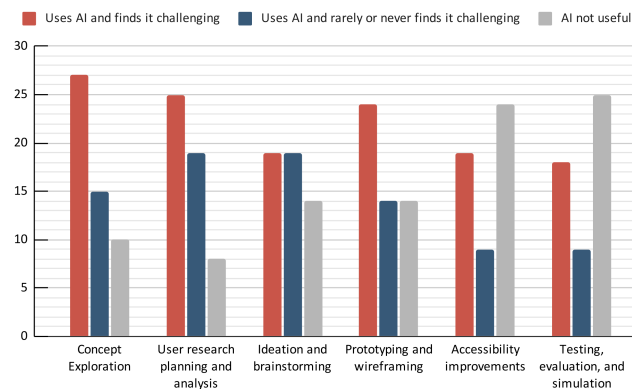


Figure 6.4: Chart illustrating the combined results of how useful and challenging designers found AI tools across six design tasks. The red bars indicate the overlap in which designers found AI tools both useful and challenging.

To sum up, the results of the quantitative data show that designers are using AI in many different parts of their work, and the level of integration was rated as at least occasional by most designers. Additionally, how helpful and problematic they find tools often overlaps.

6.1.2 Result based on thematic analysis

From the thematic analysis, a contradictory view on AI was found. Firstly, many designers mentioned its ability to speed up the process and help them ideate. However, when describing challenges, generic ideas and spending time prompting and iterating to achieve the desired result were mentioned. This aligns with the quantitative result, which presented a contradictory overlap in AI usefulness and challenges. Where things

seemed to get more complicated was in tasks that require judgment, creativity, or a deeper understanding of the user. Many participants described feeling unsure about how to get the most out of AI in these situations; they use it, but are not always happy with what they get.

An explanation for this might be AI sycophancy, bias, and hallucinations. Hallucinations were mentioned by only 2 respondents, while sycophancy was mentioned by 1 respondent. This might be a sign of a lack of knowledge about AI, which can influence how designers interact with the tool and what they expect it to be capable of.

Taken together, the survey showed that most designers have integrated AI into their work to some degree, but it is still being figured out. Designers keep returning to AI for help, believing in its potential to speed up the process, broaden perspectives, and explore alternatives. At the same time, they find it challenging to produce useful results, ultimately spending time instructing instead of creating. In the end, designers are optimistic about the tools' potential, but the gap between what they hope AI can do and what it actually delivers in practice remains significant for many.

6.2 What challenges and opportunities emerge when UX designers interact with AI in concept exploration and ideation? (RQ2)

This research question was answered through the thematic analysis of the workshop. This resulted in 7 themes describing the challenges and opportunities of using AI for concept exploration and ideation. Here, we will examine how the themes address the research question.

A recurring challenge was the lack of structure in how designers approached the task, described in *Theme 2: Prompting Quality Determines Output Quality*. However, an opportunity arose when designers became more deliberate in how they framed prompts, and the results seemed to improve noticeably.

This led to another challenge described in *Theme 5: Output Generality and the Need for Specificity*, where they easily ended up in a loop of similar-feeling results. This was mainly thought of as a problem if proper prompting had not been conducted, but some groups also discovered that output easily fell back into generic, narrowed responses over time.

A third challenge, described in *Theme 6: From Linear Chat to Visual, Spatial AI Workspace*, was the chat interface most AI tools use. Designers are used to being able to see all their ideas at once, on a whiteboard, on sticky notes, or in a tool like FigJam, so they can compare, cluster, combine, branch out, and build on them. In a chat window, earlier ideas scroll out of view, and it becomes hard to get a sense of the whole. Several

participants mentioned this as something that made the process feel less useful than it could be.

Another related challenge was that, although some of these features may exist in certain tools, finding the right tools and keeping up with what AI tools are currently capable of is itself a challenge for many designers. This was described in *Theme 3: Conflicting Understandings of AI capabilities*, which highlights designers' conflicting understanding of AI tools as a challenge. Beyond keeping up with new tools, designers seemed to perceive capabilities associated with intelligence that do not exist.

On the opportunity side, AI helps designers get started. *Theme 1: AI as a Warm-Up and Cognitive Kickstarter*, described how designers used it to reduce the blank-page feeling and quickly give them something to react to and get started. Participants who used it this way, treating it as a divergence tool rather than expecting polished ideas, generally had a better experience with it.

To sum up, designers expressed their appreciation for and need to use AI early on for inspiration, to get unstuck, or to get in the mood. However, the AI was quick to provide detailed ideas and solutions, which were at times deemed too general. In addition, the chat limited their ability to compare and combine ideas, and they felt that complementary and affirming responses weren't genuine. For a task like ideation, where the whole point is to find new solutions, this is a real limitation. Designers tried to push back against this by providing the prompt with more context and specifications, resulting in responses they found more relevant and useful. The recommendations in the next section will address these challenges.

6.3 What are recommendations that can improve the use of AI in concept exploration and ideation, to support UX designers and improve their experience? (RQ3)

The final recommendations were structured into three sections: Recommended knowledge, general recommendations and task-specific recommendations. To answer Research Question 3, the most essential extracts from each recommendation will be presented. The full version of the recommendations can be found in Appendix B.

The recommended knowledge was created to give the reader valuable information that might influence their perception and use of AI. It also served to support the recommendations with relevant research. However, in this chapter, the recommended knowledge section will be presented as

a recommendation and the information it contains will instead be cited directly cited within each relevant recommendation.

6.3.1 Recommendation Knowledge

During the workshops, a mismatch between designers' understanding of how AI works and how designers are affected by the tool was found. As a result, many of the recommendations are built on a foundation that explains AI's capabilities and limitations to utilise the tool to its full potential. Since research indicates it is more than just a neutral tool and can unconsciously negatively affect users, information about this was also deemed important to demonstrate and include in the recommendations.

For this reason, the recommended knowledge section was created to provide readers with relevant background information that may influence their perception and use of AI. This section also supports the recommendations through relevant research. However, the knowledge presented in this section can also be understood as a recommendation in itself, since it may help designers use AI more critically and effectively.

In this chapter, the recommended knowledge section is therefore presented as one recommendation. The specific information included in that section will not be described in full here. Instead, the most relevant points will be cited directly within each related recommendation, while the full recommended knowledge section can be found in Appendix B.

6.3.2 Recommendation 1: Direct the output through context

LLMs use transformers, which were developed for Google search, a model that predicts which words in a sentence are most important rather than focusing on every word in the sentence [71]. This way, context can guide the model towards more relevant and specified responses [45], [72]. However, they do not guarantee higher quality or correctness.

Research finds that AI advice often fails to adapt to context and situation, yet creates the impression of being insightful. Dergaa et al. found that using AI for medical advice resulted in inaccurate assessments, yielding generic, similar responses that failed to adapt to the different contexts and scenarios [73]. Another study from Harvard found that context had little effect on the validity of business advice [74]. Prompting different LLMs with different scenarios and amounts of context showed that similar advice was given regardless. They found that the main factor driving variation in advice was the order in which options were presented, choosing one of the top options regardless of the situation. Although the studies do not assess the validity of design solutions and ideas, they assess the ability to evaluate situations and provide advice accordingly across two distinct domains. This highlights the model's limitations in analysing, resonating,

and adapting suggestions, which may have negative impacts in ideation as well.

In addition to the response’s accuracy, context was found to make advice look more tailored and specific, even though it wasn’t [74]. Dergaa et al. noted its capability to extend empathy towards the patient [73]. As explored further in the section above, this is likely due to the sycophantic design of responses, which tricks us into thinking something is insightful when it is delivered in a personal, authoritative, and confident way [75]. Designers should therefore use context, instructions, constraints, and examples of the wanted output to direct the output, while still critically assessing the result.

6.3.3 Recommendation 2: Challenge and stay critical of AI responses

A faulty understanding of capabilities, sycophancy and heightened trust can also increase the designer’s reliance on the tool [76], [77]. It may be tempting to use AI to reduce mental demand, but not being challenged can lead to cognitive decline and the loss of important skills, such as critical thinking, judgement and decision-making [78], [79], [80], which are essential to possess as a designer. Additionally, creativity is a skill [81], that we risk losing if it is not exercised. Designers should therefore be wary of which responsibilities and forms of work they delegate, since giving away too much of the creative process may reduce both creative practice and control.

Designers should therefore actively challenge AI-generated responses instead of accepting them as reliable or objective. Since AI can present information in a confident and human-like way, it may be perceived as more valid than it is. Understanding AI’s limitations and reducing human-like or overly validating responses can help designers maintain control of the design process.

6.3.4 Recommendation 3: Change the settings to customise responses and memory

AI settings can be adjusted to make responses more concise, direct, and less validating. This can help designers avoid being overwhelmed by long answers and reduce some of the effects of sycophantic responses. Designers should also understand how memory works, since stored or shared context can influence responses in both useful and unwanted ways.

However, These customisations do not eliminate sycophantic design or hallucination. They reduce their surface expression, which lowers unearned trust, but the underlying model behaviour remains. This can cause responses to still confirm the designer’s view, but less explicitly,

making it more difficult to detect.

6.3.5 Recommendation 4: Use AI to spark inspiration while keeping creative control

AI output can further influence designers through cognitive bias and anchoring, which influence how one processes information and makes decisions [82], [83]. An early study on cognitive bias shows that even unrelated information can affect decision-making [84]. A popular example from this study illustrates how participants' guesses were influenced by the random number they obtained by spinning a wheel. Therefore, once AI is used for ideation, it may shape how the designer frames the problem, evaluates alternatives, or decides which ideas are worth developing. The designer should therefore be cautious about how and when AI is used, especially if it is applied early in the ideation process.

Therefore, using AI early on is not recommended, but if it is used as a way to warm-up or a source of inspiration, the designers should avoid letting it produce finished-looking ideas and solutions. Instead, AI should be used to generate low-fidelity triggers, such as mood boards, rough sketches, or incomplete ideas that leave space for the designer's own imagination. This helps reduce the risk of becoming too influenced by generic or biased AI output.

6.3.6 Recommendation 5: Generate diverse and rough AI ideas, then explore independently

Anthropomorphism is used to make the tool more user-friendly and to heighten users' trust in it [85], [86]. This has been shown to reduce users' likelihood of questioning AI advice [86]. Sycophancy increases this risk by making responses feel valid, reliable, or contextually appropriate [75], [76], [87]. When a system is overly agreeable, affirming, or confidence-inducing, it can reduce the designer's critical distance from the output [76]. Additionally, studies show that this feature can create an echo chamber in which the user's ideas are amplified rather than critically questioned [77], [88], [89]. This creates a delusion, and in the worst case, psychosis. As the output contains bias, the types of answers received are also limited by the training data and by how RLHF has shaped the model. As illustrated in Figure B.5, this can limit the scope of ideas explored by the designer while also having a feeling of shaping ideas and being in control. For these reasons, questioning output and being aware of how one is influenced can be challenging and may affect the development of ideas.

Therefore, AI should preferably be used after designers have first explored their own ideas. When using AI for ideation, designers should ask for many varied and rough ideas, explore the less obvious ones first, and move the ideas outside the chat into a canvas space. This supports

broader exploration while keeping selection and creative development in the designer's hands.

6.3.7 Recommendation 6: Do not use AI for advice, use it to challenge your ideas

Anthropomorphism has also been shown to heighten users' trust in AI advice, which also reduces critical thinking [86]. As a result, the content may appear adapted and, at times, even be valid. However, the relevance and quality of its suggestions and advice cannot be reliably trusted, and the designer should therefore always reflect upon the output and stay critical of it.

Therefore, AI should not be used to select the best idea or make design decisions, since its suggestions can be biased and influenced by the order of presented options. Instead, one can use AI to ask counter-questions about the design, prompting the designer to reflect and improve it using their own knowledge and judgement. However, be aware that the responses, even when challenging, are still tailored to please the user. In the end, the tool merely imitates how advice, reasoning, and empathy are expressed and cannot adapt advice to context due to the absence of true understanding. Starting a new chat for each idea can also help avoid unwanted influence from previous context.

7

Discussion

In this chapter we will first discuss how well the research questions were answered, and potential limitations. After that, additional risks tied to AI usage will be explored, and lastly we will delve into ethical considerations that were not included in the final recommendations. Through this, we will uncover some interesting findings that didn't make it to the final recommendations.

7.1 Limitations of the Final Recommendations

The recommendations that came out of this study, of course have some limitations that should be acknowledged. Due to the qualitative and exploratory nature of the study, it has primarily focused on identifying potential needs and solutions, with less attention to refinement and validation.

The recommendations should therefore undergo further testing, not only through evaluation but also in practice, to assess their performance in concept exploration and ideation. How readers react to the information section should also be evaluated to understand how it does or does not inform and affect usage, and how it can be adapted for easy digestion and application.

Since much of the research that grounded the recommendations in the information section drew on studies outside our main domain, this introduced the potential for faulty understanding and misuse. Therefore, in addition to users, the recommendations and information should be evaluated by domain experts in AI, psychology, and neuroscience to ensure they align with their expertise.

The recommendations themselves also have limitations, as not all feedback was taken into consideration due to time constraints. For example, the participants requested more figures and examples of how to apply the recommendations in a specific scenario. Adding visual elements would have been a valuable addition to the recommendations, hopefully making them lighter and easier to grasp. Additionally, the workshops involved

a small number of participants, and the findings are not intended to generalise broadly. AI tools are also evolving fast, which means some observations may date quickly.

Still, the broader issue this thesis points to is unlikely to disappear, since the core of how LLMs are built and function has not advanced much beyond their ability to mimic language. Research highlights AI’s potential to negatively influence users’ cognitive biases by introducing generic, biased ideas [83]. It is also good at convincing the user that the output is relevant and specified through anthropomorphism and sycophancy, even when it is not [76], [86]. This makes the advice look tailored to the situation, but it lacks depth, and true alignment and understanding of the situation, regardless of context [73], [74].

Additionally, this can lead to AI reliance and cognitive decline, resulting in the loss of important skills such as critical thinking and creativity [78], [80]. If many designers use the same AI tools in the same ways during early-stage creative work, the field’s collective creative output risks becoming more homogeneous over time. That is worth taking seriously. Better practices, clearer awareness of AI’s limitations, and a stronger emphasis on keeping human judgment central are all steps in the right direction.

7.2 Risks of using AI in design

So far, we have discussed our findings and what stood out. We will now explore how well the last research question, which produced the recommendations, accounts for risks tied to AI usage, and identify other risks that might influence the designer and warrant further investigation.

7.2.1 Risk of deceptive marketing

However, there is also a deceptive element to LLM, especially in its marketing, that we want to diver deeper into. LLMs can be seen as deceptive, both in its design and in its marketing. Before ChatGPT was first released, OpenAI described the technology as potentially dangerous and suggested that it was not ready for public release. Yet only a few months later, it was made freely available. At the same time, AI has often been framed as a step toward AGI and as a technology capable of solving major global problems, provided it is developed by the “right” people. Both extreme claims exaggerate its current capabilities and influence what users, companies, and UX designers believe the technology can do.

7.2.1.1 Ending or saving the world?

Karen Hao, an investigative journalist with a background in software, has followed the development of AI since 2019 and writes of the journey

in her book "The Empire of AI". She discovered a misalignment between the capabilities their systems could perform and the how it was marketed, their outspoken goals and how it was followed up on, as well as hidden ethical dilemmas that she found important to bring to the surface. [71]. One of the early claims she reacted on was that the technology had the potential to destroy the world if released to early or developed by the wrong people. Not long after the first version of chatGPT was released despite their statement. Due to the misalignment in claims and action they should be questioned critically. Why would a company describe its own product as potentially world-ending while still releasing it to the public?

At the same time, openAI claims to be on the path of developing AGI according to their own website [90]. They have also been making claims that developing their products is essential to solve the world problems we are faced with, and although the energy consumption of data centres may have a significant cost on the environment at the moment, this is also what will save the world [71]. However, these claims are not supported by sufficient evidence. So far, AI has had significant effects on society, but we have yet to see proof that it is solving major global problems.

However, they are not alone in thinking the technology has the potential to do good, as the book AI for greater good explains LLMs potential to solve world problems [72]. However, although the book explains the differences between types of AI, such as LLMs, expert systems, and narrow AI, it seems to conflate their respective capabilities and achievements. For example, AI is described as a broadly successful technology, but the discussion appears to place LLMs on a pedestal by drawing on the successes of other, largely unrelated AI systems.

ultimately, this narratives creates a false sense of what LLMs are capable of. Presenting the technology as being on an inevitable path toward AGI risks overstating its abilities and encourages designers and users to trust the system beyond what it can reliably deliver. Additionally, focusing on future catastrophic risks can shift attention away from the harms AI is already causing and give companies a reason to continue development in the name of safety and control. The technology may contribute to serious harm, but not necessarily because it is as intelligent or powerful as these claims suggest.

7.2.1.2 The design of the tool

Beyond how LLMs are marketed, the design of the tool further increases the deception. As the information chapter of the recommendations describes, mimicing a chat interface can make it easy for the user to understand how the tool should be used. However, drawing on their existing mental model of messaging platforms, this design also shapes how

users interpret the system’s capabilities. When waiting for a response, the interface often resembles a human typing, which can make the system feel more intelligent than it actually is. Status messages such as “thinking” are also misleading, since they suggest human-like reasoning when what it actually does is statistical text generation.

Although the chat format makes the tool accessible, it also encourages users to anthropomorphise the tool, and treat it as if it understands, reflects, or communicates like a person. This can lead users to overestimate the system’s reliability, agency, and competence. This may influence UX designers to have a faulty perception of what the system is capable of, and delegate tasks to the system that might not produce sufficient outcomes. Due to the responses being sycophantic insufficient results may also be difficult to detect. These impacts are discussed in greater detail in the information section, and hope to inform the designer in order to avoid giving their power away and risking to lose skills.

7.2.2 Impact on Designers’ mental health and quality of product

There are many ways that designers’ mental health can be negatively affected when using LLMs. As described above in the tool’s design, the chat format poses a risk that users develop a human-like perception of the tool. This is because the brain doesn’t fully comprehend that there isn’t a human on the other side. This imposes additional risks for the designer’s mental health, such as delusion, psychosis and cognitive decline.

The perception of human-like features and intelligence can also lead to a faulty mental model of the tool. Designers may assume that the system has capabilities it does not actually have, which affects how they use it and what tasks they rely on it for. This can make designers less likely to question the output critically, increasing the risk that errors, assumptions, or biases are carried into the design process. As a result, while LLMs may speed up parts of the design process, this efficiency can come at a cost if the designer is not aware of what the tool is capable of and how it may influence them. In the end, if this hasn’t been considered, it risks affecting the product and the end user.

Due to this, we made sure to include information of how the LLM’s can influence the design and how they may impact the designer. By making these risks visible, the recommendations encourage designers to use the tool more critically and to reflect on when its use is appropriate, rather than treating it as a neutral or fully reliable part of the process.

However, the risk of delusion and psychosis was only briefly mentioned, and not tied to how it could affect the designer beyond the development of their ideas. Dalmasso et al. suggest that the beginning of harmful use can often begin with a mundane task, which could potentially be

job-related [87]. When met with empathy and understanding, the user may return later for support in other, more personal areas of their life. They may be a potential risk to the designer if using AI often and getting increasing feelings of being heard and understood, which may be worth including in recommendations as well.

7.3 Ethical considerations

The recommendations have focused mainly on how the designer and their ideas are affected, but this section will explore the outer consequences of AI usage and its ethical implications. This includes people affected by AI use, such as creators of data used to build AI models, workers in the AI industry, end users, and people living near data centres.

7.3.1 Bias, cultural values and stereotypes

AI systems contain biases, cultural values, and stereotypes related to gender, ethnicity, age, class, disability, and culture. As explained before, AI-generated outputs unconsciously influence our minds, creating and amplifying cognitive bias. This might not only affect the user but also cause a ripple effect and infiltrate the societal worldview.

Fixing this is difficult because of how the neural network is built and its need for a vast amount of information. In the early versions of ChatGPT, data was filtered, but as a result, although the system contained cleaner data, it also performed considerably worse due to the smaller dataset. The act of filtering data was later abandoned since it was deemed too time-consuming.

After reinforcement learning, the model also learns from human feedback (RLHF), also called “behaviour training”. During this training, people working as data labellers adjust what the output should and should not be. This introduces additional bias, cultural values, and stereotypes. It also raises the question of which behaviours are being enforced and according to whose values. With long conversations, these preferences can be broken down, called jailbreaking, even though it is not the original type of jailbreak. However, this could possibly create even more harmful outputs.

7.3.2 Data labellers

Working with unfiltered data also affects the data labellers who later have to review harmful AI-generated content. Karen Hao investigated this issue and interviewed data labellers to understand how the work affected them [71]. They described being badly affected psychologically by the abusive language and disturbing images they were required to process. Through this, she also uncovered a toxic and exploitative system in which

people in lower-income countries with vulnerable economic situations were targeted.

Because of the workers' vulnerable circumstances, they were left waiting for assignments with little stability or protection, often putting their lives on hold in the hope of receiving paid work [71]. Many feared missing an opportunity, since a single task could determine whether they could afford basic necessities such as food. Over time, this uncertainty began to control their daily lives, including their sleep, with some workers waking several times during the night or staying awake to avoid missing an opportunity.

Based on this investigation, it is clear that data labellers are being exploited and harmed, and that the use and promotion of LLMs can contribute to continued harm. Creating recommendations indirectly promotes the use of the tool and should therefore take this into consideration. The effect on those who make the tool usable is rarely discussed. However, this should be highlighted more to better understand how designers' use of AI and our recommendations, which promote its use, affect others.

7.3.3 Data centres

It is not only data labellers who are negatively affected by the development of AI, but also people living near the areas where data centres are built. Large data centres are required to host these systems, and they can generate constant noise that disrupts the local sound environment. This creates a continuous background hum that can affect residents' quality of life, sleep, and overall well-being [71].

7.3.4 Crediting artists

The work of others is being used without acknowledgement and economic compensation. Because the system cannot trace the source of information, it is also not possible to appropriately credit the creator whose work the output is based on. Even if it were possible to credit the artist, one might wonder whether using others' work without any economic compensation is fair. This is an issue that the recommendations do not yet adequately address or provide advice on.

The reason this is legally possible is partly that the technology is outpacing legislation. The tool also avoids responsibility for aligning with CC policy by claiming that the system itself doesn't replicate anyone's work and that its output cannot be controlled. It is therefore up to the user to ensure they do not replicate anyone's work, a risk the recommendations do not warn designers about.

8

Conclusion

This study aimed at understanding how UX designers use AI across the design process, where the biggest challenges and opportunities lie, and how to form recommendations accordingly. This was done by first collecting data through an online survey, which informed our focus, concept exploration and ideation. A workshop was then conducted with 11 UX designers to explore how they used AI for concept exploration and ideation, their mental model of the system, and what they found useful and challenging. This was analysed using thematic analysis to develop themes of participants' needs and usage, a workflow of how AI tools were used, and a map of how designers resonated about their AI use in different stages. The analysed data then helped inform and develop the first iteration of recommendations, together with research on AI capabilities and their impacts on users. The recommendations were reviewed in a follow-up focus group with six of the 11 workshop participants. This generated feedback on how the recommendations were received, focusing on their clarity, applicability, and relevance. Based on this, the recommendations were adapted according to the feedback.

These findings shaped the six recommendations that form the core contribution of this thesis. Thanks to the workshop, we could identify the needs tied to AI usage in concept exploration and ideation, what they hoped to achieve, and where their knowledge of AI could lead to a faulty understanding of its capabilities. Combined with research about AI capabilities and user impacts, recommendations were developed to address designers' underlying needs rather than how they expressed that they wanted to use it. Given the identified mismatch between designers' understanding of AI and its actual capabilities, a stronger understanding of the technology could help designers use the tool more intentionally. Therefore, a section on AI and its potential influence on the designer was added to the recommendations.

All in all, our thesis showed a mismatch between designers' understanding of AI chatbots and what they can actually do. Designers showed a desire to use AI chatbots to get started, generate a wide range of alternative ideas and develop the chosen ones together with AI. They also expressed the importance of proper prompting to obtain valid outputs, which

included detailed heavy context and a specific scenario. The designers' needs were taken into consideration when developing recommendations, and were adapted based on how research suggests that designers can be negatively impacted. Together, these findings highlight the need for clearer guidance on how AI chatbots can support designers without limiting their creativity, judgment, or role in the design process.

Bibliography

- [1] Y. Luo, “Designing with ai: A systematic literature review on the use, development, and perception of ai-enabled ux design tools,” *Advances in Human-Computer Interaction*, vol. 2025, no. 1, p. 3869207, 2025.
- [2] N. P. Utama, P. E. L. Tobing, and Y. Kurniawan, “The influence of artificial intelligence on user experience design: Systematic literature review,” in *2025 IEEE Symposium on Industrial Electronics & Applications (ISIEA)*, IEEE, 2025, pp. 1–6.
- [3] K. Navaneeth Sankar, T. Singh, K. Sharma, and A. K. Sinha, “On the evolution of user experience design, with artificial intelligence and machine learning,” in *International Conference on Research into Design*, Springer, 2025, pp. 243–252.
- [4] M. Muehlhaus and J. Steimle, “Interaction design with generative ai: An empirical study of emerging strategies across the four phases of design,” *arXiv preprint arXiv:2411.02662*, 2024.
- [5] Y. Lu, Y. Yang, Q. Zhao, C. Zhang, and T. J.-J. Li, “Ai assistance for ux: A literature review through human-centered ai,” *arXiv preprint arXiv:2402.06089*, 2024.
- [6] S. Wadinambiarachchi, R. M. Kelly, S. Pareek, Q. Zhou, and E. Velloso, “The effects of generative ai on design fixation and divergent thinking,” in *Proceedings of the 2024 CHI conference on human factors in computing systems*, 2024, pp. 1–18.
- [7] M. Takaffoli, S. Li, and V. Mäkelä, “Generative ai in user experience design and research: How do ux practitioners, teams, and companies use genai in industry?” In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, 2024, pp. 1579–1593.
- [8] Z. Zhu, H. Lee, Y. Pan, and P. Cai, “Ai assistance in enterprise ux design workflows: Enhancing design brief creation for designers,” *Frontiers in artificial intelligence*, vol. 7, p. 1404647, 2024.
- [9] J. Li, H. Cao, L. Lin, Y. Hou, R. Zhu, and A. El Ali, “User experience design professionals’ perceptions of generative artificial intelligence,” in *Proceedings of the 2024 CHI conference on human factors in computing systems*, 2024, pp. 1–18.
- [10] J. Tholander and M. Jonsson, “Design ideation with ai-sketching, thinking and talking with generative machine learning models,” in

- Proceedings of the 2023 ACM designing interactive systems conference*, 2023, pp. 1930–1940.
- [11] S. Kapania, W. Agnew, M. Eslami, H. Heidari, and S. E. Fox, “Simulacrum of stories: Examining large language models as qualitative research participants,” in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, pp. 1–17.
 - [12] H. Schroeder, M. A. L. Quéré, C. Randazzo, D. Mimno, and S. Schoenebeck, “Large language models in qualitative research: Can we do the data justice,” *arXiv preprint arXiv:2410.07362*, 2024.
 - [13] E. Kirsten, A. Buckmann, L. Lassak, N. Borgert, A. Mhaidli, and S. Becker, “From assistance to autonomy—a researcher study on the potential of ai support for qualitative data analysis,” *arXiv preprint arXiv:2501.19275*, 2025.
 - [14] E. Kuang, M. Li, M. Fan, and K. Shinohara, “Enhancing ux evaluation through collaboration with conversational ai assistants: Effects of proactive dialogue and timing,” in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–16.
 - [15] C. Akverdi and G. E. Baykal, “Generative ai tools in design fields: Opportunities and challenges in the ideation process,” in *Adjunct Proceedings of the 2024 Nordic Conference on Human-Computer Interaction*, 2024, pp. 1–5.
 - [16] W.-F. Wang, C.-T. Lu, N. Ponsa i Campanyà, B.-Y. Chen, and M. Y. Chen, “Aideation: Designing a human-ai collaborative ideation system for concept designers,” in *Proceedings of the 2025 chi conference on human factors in computing systems*, 2025, pp. 1–28.
 - [17] H. Shen, L. Shen, W. Wu, and K. Zhang, “Ideationweb: Tracking the evolution of design ideas in human-ai co-creation,” in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, pp. 1–19.
 - [18] T. Kneareem, M. Khwaja, Y. Gao, F. Bentley, and C. E. Kliman-Silver, “Exploring the future of design tooling: The role of artificial intelligence in tools for user experience professionals,” in *Extended Abstracts of the 2023 CHI conference on human factors in computing systems*, 2023, pp. 1–6.
 - [19] M. Jacobs, S. Amershi, C. J. Cai, J. Wortman Vaughan, H. Wallach, and D. S. Weld, “Overreliance on AI: Literature review,” Microsoft Research, Tech. Rep., 2022. [Online]. Available: <https://www.microsoft.com/en-us/research/wp-content/uploads/2022/06/Aether-Overreliance-on-AI-Review-Final-6.21.22.pdf>.
 - [20] Base9, *Cognitive offloading & AI: Risks, bias and control*, Accessed: 2026-05-27, 2026. [Online]. Available: <https://www.base9.it/en/blog/21/working-with-artificial-intelligence-and-the-threat-of-cognitive-offloading.html>.

-
- [21] K. B. Q. Le and W. H. Kunz, “When humans stop thinking: Tackling the silent threat of ai complacency in service operations,” *Journal of Service Management*, vol. 37, no. 6, pp. 78–118, 2026.
 - [22] B. M. Chaudhry, “Concerns and challenges of ai tools in the ui/ux design process: A cross-sectional survey,” in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–6.
 - [23] H. Subramonyam, D. Thakkar, A. Ku, J. Dieber, and A. K. Sinha, “Prototyping with prompts: Emerging approaches and challenges in generative ai design for collaborative software teams,” in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, pp. 1–22.
 - [24] D. Norman and J. Nielsen, *The definition of user experience (ux)*, Accessed: 2025-02-12, 1998. [Online]. Available: <https://www.nngroup.com/articles/definition-user-experience/>.
 - [25] D. Winter, “Will the design of the human–product relationship follow user experience?” *i-com*, vol. 23, no. 2, pp. 133–138, 2024.
 - [26] K. Holtzblatt and S. Holtzblatt, “Communicating user research in order to drive design and product decisions,” in *CHI’14 Extended Abstracts on Human Factors in Computing Systems*, 2014, pp. 1155–1158.
 - [27] N. BEUDON, “Ux design in the library,” *Marketing for Libraries and Information Services: Systems and Developments*, pp. 13–23, 2025.
 - [28] A. Shokrizadeh, B. Bahati Tadjuidje, S. Kumar, S. Kamble, and J. Cheng, “Dancing with chains: Ideating under constraints with uidec in ui/ux design,” in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, pp. 1–23.
 - [29] J. Innes, “Defining the user experience function: Innovation through organizational design,” *interactions*, vol. 14, no. 3, pp. 36–37, 2007.
 - [30] G. Corrêa, R. Pereira, M. S. Silveira, and I. Gasparini, “Exploring usability and user experience evaluation methods: A tertiary study,” in *ICEIS (2)*, 2024, pp. 357–368.
 - [31] S. Kremer, S. Günther, and U. Lindemann, “Roles in user experience design-transferring insights from experience oriented disciplines,” in *Proceedings of the 10th International Conference on Design & Emotion, Celebration & Contemplation, 27-30 Sept. 2016, Amsterdam*, 2016, pp. 473–482.
 - [32] B. Jouda, N. Owies, and L. F. Al-Qora’n, “The role of user experience in software sustainability designing for longevity,” in *2025 16th International Conference on Information and Communication Systems (ICICS)*, IEEE, 2025, pp. 1–6.
 - [33] H. W. J. Rittel and M. M. Webber, “Dilemmas in a general theory of planning,” *Policy Sciences*, vol. 4, pp. 155–169, 1973. DOI: <https://doi.org/10.1007/BF01405730>.

- [34] J. Falk, *AI för nybörjare 2.0: Allt du behöver för att komma i gång med artificiell intelligens*. Natur Kultur Läromedel och Akademi, 2025.
- [35] J.-Y. Mao, K. Vredenburg, P. W. Smith, and T. Carey, “The state of user-centered design practice,” *Communications of the ACM*, vol. 48, no. 3, pp. 105–109, 2005.
- [36] G. Corrêa, R. Pereira, M. S. Silveira, and I. Gasparini, “Exploring usability and user experience evaluation methods: A tertiary study,” vol. 2, 2024, pp. 357–368. DOI: 10.5220/0012606100003690.
- [37] D. Yang, Y. Sun, D. di Stefano, and M. Turrin, “A computational design exploration platform supporting the formulation of design concepts,” in *2017 Proceedings of the Symposium on Simulation for Architecture & Urban Design*, vol. 8, 2017.
- [38] P. Sarkar, A. Chakrabarti, et al., “Understanding search in design,” in *DS 42: Proceedings of ICED 2007, the 16th International Conference on Engineering Design, Paris, France, 28.-31.07. 2007*, 2007, pp. 319–320.
- [39] T. Brown et al., “Design thinking,” *Harvard business review*, vol. 86, no. 6, p. 84, 2008.
- [40] A. M. Webb, “Pen+ touch diagramming to stimulate design ideation,” in *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, 2015, pp. 331–332.
- [41] P. Gärdenfors, *Kan AI tänka? Om människor, djur och robotar*. Fri tanke förlag, 2025.
- [42] W. Liu, G. Zhuang, X. Liu, S. Hu, R. He, and Y. Wang, “How do we move towards true artificial intelligence,” in *2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)*, IEEE, 2021, pp. 2156–2158.
- [43] D. McMeel, S. Jayachandran, and C. Walker, *Will humans replace ai? a design dilemma*, 2024.
- [44] M. S. Graziano, *Consciousness and the social brain*. Oxford University Press, 2013.
- [45] S. Wolfram, *What Is ChatGPT Doing:... and Why Does It Work?* Wolfram Media, 2023.
- [46] S.-W. Hsiao and Y.-J. Tsai, “Application of multi-diamond processing model to hairstyling,” *Engineering Proceedings*, vol. 74, no. 1, p. 21, 2024.
- [47] S. Sunsunwal and P. Agarwal, *Double diamond design process*, https://informatics.nic.in/uploads/pdfs/e9eed640_34_35_tu_double.pdf, Accessed: 2025-02-15, 2023.
- [48] J. W. Knopf, “Doing a literature review,” *PS: Political Science & Politics*, vol. 39, no. 1, pp. 127–132, 2006.

-
- [49] H. Sharp, P. Jenny, and Y. Rogers, "Interaction design:: Beyond human-computer interaction," 2007.
 - [50] S. A. Perrig, L. F. Aeschbach, N. Scharowski, N. von Felten, K. Opwis, and F. Brühlmann, "Measurement practices in user experience (ux) research: A systematic quantitative literature review," *Frontiers in Computer Science*, vol. 6, p. 1368860, 2024.
 - [51] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative research in psychology*, vol. 3, no. 2, pp. 77–101, 2006.
 - [52] S. K. Ahmed et al., "Using thematic analysis in qualitative research," *Journal of Medicine, Surgery, and Public Health*, vol. 6, p. 100198, 2025.
 - [53] G. Harboe and E. M. Huang, "Real-world affinity diagramming practices: Bridging the paper-digital gap," in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 2015, pp. 95–104.
 - [54] C. Plain, "Build an affinity for kj method," *Quality Progress*, vol. 40, no. 3, p. 88, 2007.
 - [55] P. Storvang, B. Mortensen, and A. H. Clarke, "Using workshops in business research: A framework to diagnose, plan, facilitate and analyze workshops," in *Collaborative research design: Working with business for meaningful findings*, Springer, 2017, pp. 155–174.
 - [56] P. Storvang, A. H. Clarke, and B. Mortensen, "Workshops as a research method in business research," in *Collaborative research design: Working with business for meaningful results*, Springer, 2024, pp. 121–143.
 - [57] A. Blythe LaGasse, "Pilot and feasibility studies: Application in music therapy research," *Journal of Music Therapy*, vol. 50, no. 4, pp. 304–320, 2013.
 - [58] J. In, "Introduction of a pilot study," *Korean journal of anesthesiology*, vol. 70, no. 6, p. 601, 2017.
 - [59] A. C. Leon, L. L. Davis, and H. C. Kraemer, "The role and interpretation of pilot studies in clinical research," *Journal of psychiatric research*, vol. 45, no. 5, pp. 626–629, 2011.
 - [60] J. Wray, U. Archibong, and S. Walton, "Why undertake a pilot in a qualitative phd study? lessons learned to promote success," *Nurse researcher*, vol. 24, no. 3, 2017.
 - [61] J. Nielsen, T. Clemmensen, and C. Yssing, "Getting access to what goes on in people's heads? reflections on the think-aloud technique," in *Proceedings of the second Nordic conference on Human-computer interaction*, 2002, pp. 101–110.
 - [62] M. Fan, S. Shi, and K. N. Truong, "Practices and challenges of using think-aloud protocols in industry: An international survey," *Journal of Usability Studies*, vol. 15, no. 2, 2020.
 - [63] N. A. Jones, H. Ross, T. Lynam, P. Perez, and A. Leitch, "Mental models: An interdisciplinary synthesis of theory and methods," *Ecology and society*, vol. 16, no. 1, 2011.

- [64] B. Hanington and B. Martin, *The pocket universal methods of design: 100 ways to research complex problems, develop innovative ideas and design effective solutions*. Rockport, 2017.
- [65] S. Gibbons, *Cognitive maps, mind maps, and concept maps: Definitions*, Accessed: 2026-10-06, 2019. [Online]. Available: <https://www.nngroup.com/articles/cognitive-mind-concept/>.
- [66] K. Berezina, O. Ciftci, M. Talebi, and R. Mahdavi, “Rethinking the love-breakup letter: A three-pronged approach to understanding consumer relationships,” in *ENTER e-Tourism Conference*, Springer, 2025, pp. 461–471.
- [67] O. Doody, E. Slevin, and L. Taggart, “Focus group interviews in nursing research: Part 1,” *British Journal of Nursing*, vol. 22, no. 1, pp. 16–19, 2013.
- [68] H. Sharp, P. Jenny, and Y. Rogers, *Interaction design: beyond human-computer interaction/Rogers, Sharp, Preece*, (Sixth edition). Wiley, 2023.
- [69] ATLAS.ti Scientific Software Development GmbH, *Atlas.ti*, 2023. [Online]. Available: <https://atlasti.com>.
- [70] Nielsen Norman Group, “Care: Structure for crafting ai prompts,” May 2024, Accessed: 2026-06-16. [Online]. Available: <https://www.nngroup.com/articles/careful-prompts/>.
- [71] K. Hao, *Empire of AI: Inside the reckless race for total domination*. Random House, 2025.
- [72] R. Dodhia, *AI for social good: Using artificial intelligence to save the world*. John Wiley & Sons, 2024.
- [73] I. Dergaa et al., “Chatgpt is not ready yet for use in providing mental health assessment and interventions,” *Frontiers in Psychiatry*, vol. 14, p. 1 277 756, 2024.
- [74] A. Romasanta, D. W. Thomas Llewellyn, and N. Levina, *Researchers asked llms for strategic advice. they got “trendslop” in return*. 2026.
- [75] R. M. Batista and T. L. Griffiths, “A rational analysis of the effects of sycophantic ai,” *arXiv preprint arXiv:2602.14270*, 2026.
- [76] M. Cheng, C. Lee, P. Khadpe, S. Yu, D. Han, and D. Jurafsky, “Sycophantic ai decreases prosocial intentions and promotes dependence,” *Science*, vol. 391, no. 6792, eaec8352, 2026.
- [77] J. Huang, D. Leong, S. Chong, and M. Subramaniam, “Conversational ai and psychosis: A technological folie à deux,” *Annals of the Academy of Medicine, Singapore*, vol. 55, no. 3, pp. 163–166, 2026.
- [78] M. Gerlich, “Ai tools in society: Impacts on cognitive offloading and the future of critical thinking,” *Societies*, vol. 15, no. 1, p. 6, 2025.
- [79] M. Stadler, M. Bannert, and M. Sailer, “Cognitive ease at a cost: Llms reduce mental effort but compromise depth in student scientific inquiry,” *Computers in Human Behavior*, vol. 160, p. 108 386, 2024.

- [80] N. Kosmyna et al., “Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task,” *arXiv preprint arXiv:2506.08872*, vol. 4, 2025.
- [81] L. Lazar, “The cognitive neuroscience of design creativity,” *Journal of experimental neuroscience*, vol. 12, p. 1179069518809664, 2018.
- [82] S. Da Silva, R. Gupta, and D. Monzani, *Highlights in psychology: Cognitive bias*, 2023.
- [83] W. G. Bodamer, S. Schoenmakers, E. Nuñez Castellar, and W. A. IJsselsteijn, “Users’ mental models within human–ai interaction: A systematic scoping review,” *AI & SOCIETY*, pp. 1–18, 2026.
- [84] A. Tversky and D. Kahneman, “Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty,” *science*, vol. 185, no. 4157, pp. 1124–1131, 1974.
- [85] A. Chinmulgund, R. Khatwani, P. Tapas, P. Shah, and R. Sekhar, “Anthropomorphism of ai based chatbots by users during communication,” in *2023 3rd International Conference on Intelligent Technologies (CONIT)*, IEEE, 2023, pp. 1–6.
- [86] S. Li, Q. He, Y. Qi, and O. Turel, “Is anthropomorphizing ai worth it? the role of hands-on experience in shaping advice uptake,” *Computers in Human Behavior*, p. 108983, 2026.
- [87] J. Au Yeung, J. Dalmasso, L. Foschini, R. Dobson, and Z. Kraljevic, *The psychogenic machine: Simulating ai psychosis, delusion reinforcement and harm enablement in large language models. arxiv*, 2025.
- [88] S. Dohnány et al., “Technological folie à deux: Feedback loops between ai chatbots and mental health,” *Nature Mental Health*, pp. 1–10, 2026.
- [89] L. Osler, “Hallucinating with ai: Distributed delusions and “ai psychosis”,” *Philosophy & Technology*, vol. 39, no. 1, p. 30, 2026.
- [90] OpenAI, *Introducing chatgpt*, 2022.
- [91] J. Weizenbaum, “Eliza—a computer program for the study of natural language communication between man and machine,” *Communications of the ACM*, vol. 9, no. 1, pp. 36–45, 1966.
- [92] L. Fernandez, “Endangered judgment: Joseph weizenbaum, artificial intelligence, and the imperialism of instrumental reason,” *Critical AI*, vol. 3, no. 1, 2025.
- [93] C. Floyd, “From joseph weizenbaum to chatgpt: Critical encounters with dazzling ai technology,” *Weizenbaum Journal of the Digital Society*, vol. 3, no. 3, 2023.
- [94] M. Hedlund, *Gärdenfors, Peter, 2024. Kan AI tänka? Om människor, djur och robotar. Stockholm: Fri tanke.*
- [95] A. Wei, N. Haghtalab, and J. Steinhardt, “Jailbroken: How does llm safety training fail?” *Advances in neural information processing systems*, vol. 36, pp. 80079–80110, 2023.

- [96] S. Kushwah and N. Dave, “Ai hallucination and strategies to overcome: Enhancing human-ai interaction,” in *2025 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, IEEE, 2025, pp. 1–6.
- [97] M. T. Hicks, J. Humphries, and J. Slater, “Chatgpt is bullshit,” *Ethics and information technology*, vol. 26, no. 2, pp. 1–10, 2024.
- [98] OpenAI, *Introducing chatgpt*, 2022.
- [99] A. Alabed, A. Javornik, and D. Gregory-Smith, “Ai anthropomorphism and its effect on users’ self-congruence and self-ai integration: A theoretical framework and research agenda,” *Technological Forecasting and Social Change*, vol. 182, p. 121 786, 2022.

A

Appendix 1

A.1 Mind Maps

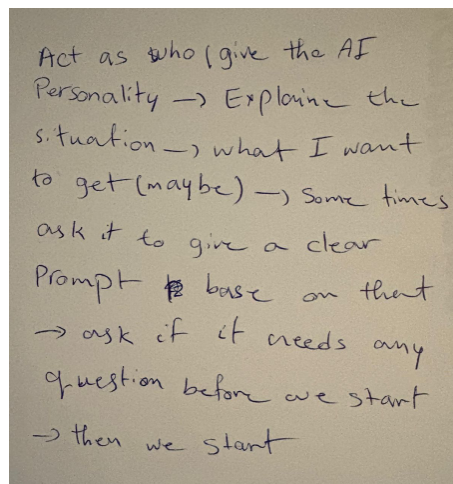


Figure A.1: Mind map: Participant 1

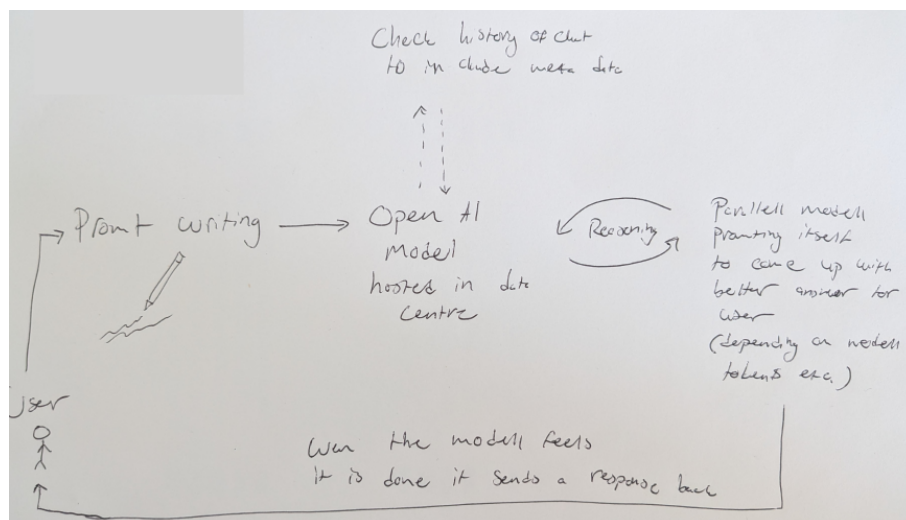


Figure A.2: Mind map: Participant 2

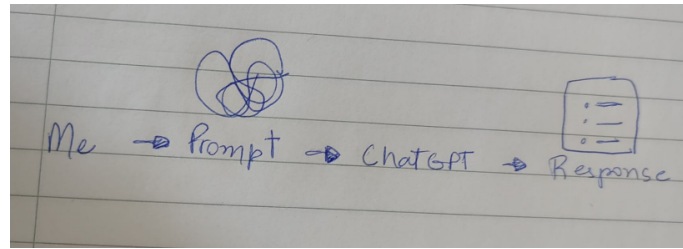


Figure A.3: Mind map: Participant 3

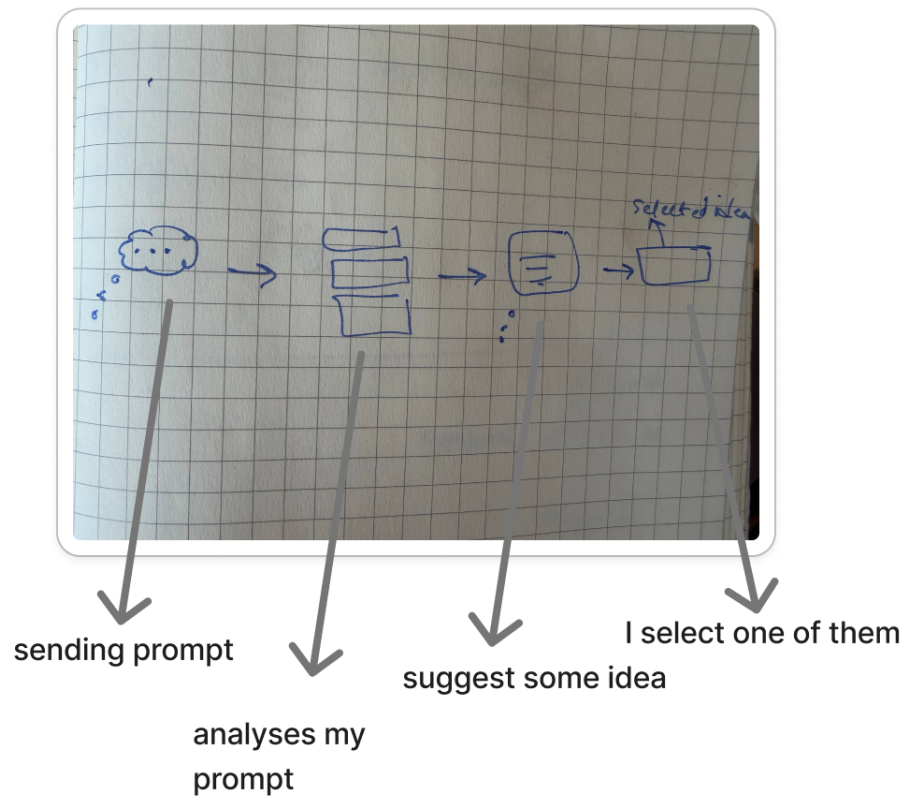


Figure A.4: Mind map: Participant 4

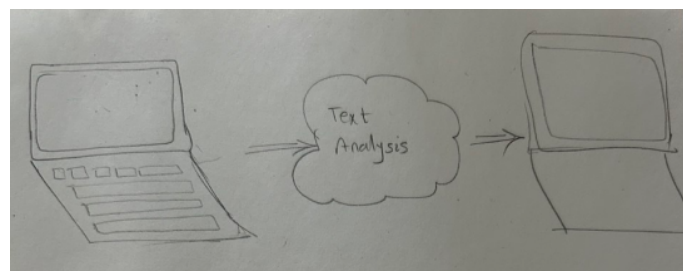


Figure A.5: Mind map: Participant 5

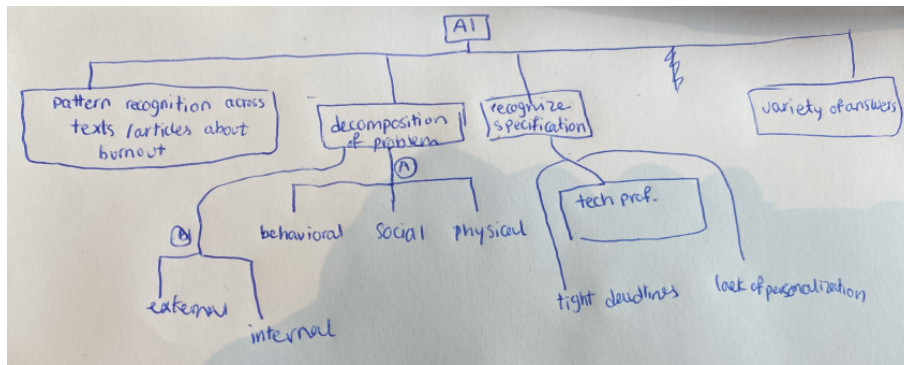


Figure A.6: Mind map: Participant 6

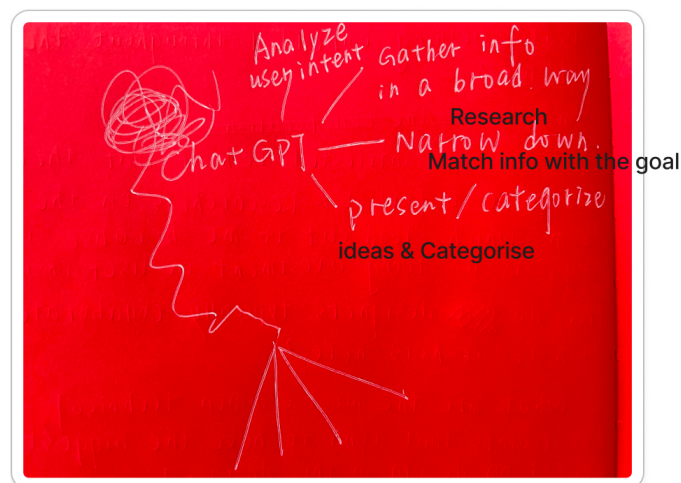


Figure A.7: Mind map: Participant 7

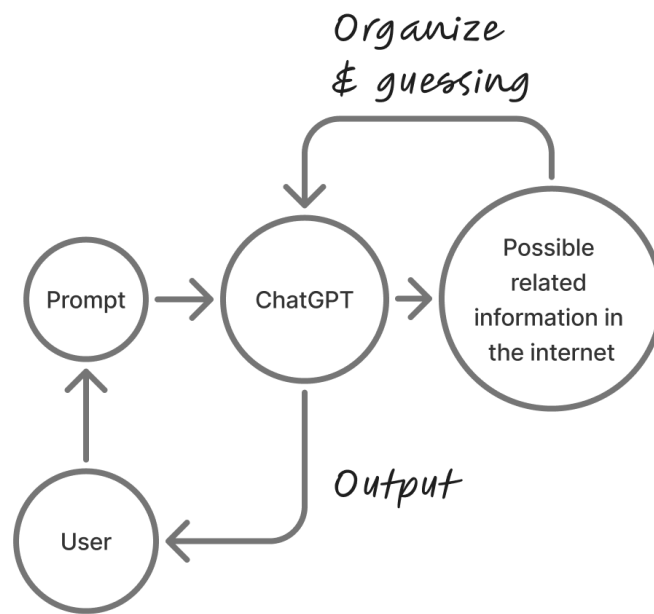


Figure A.8: Mind map: Participant 8

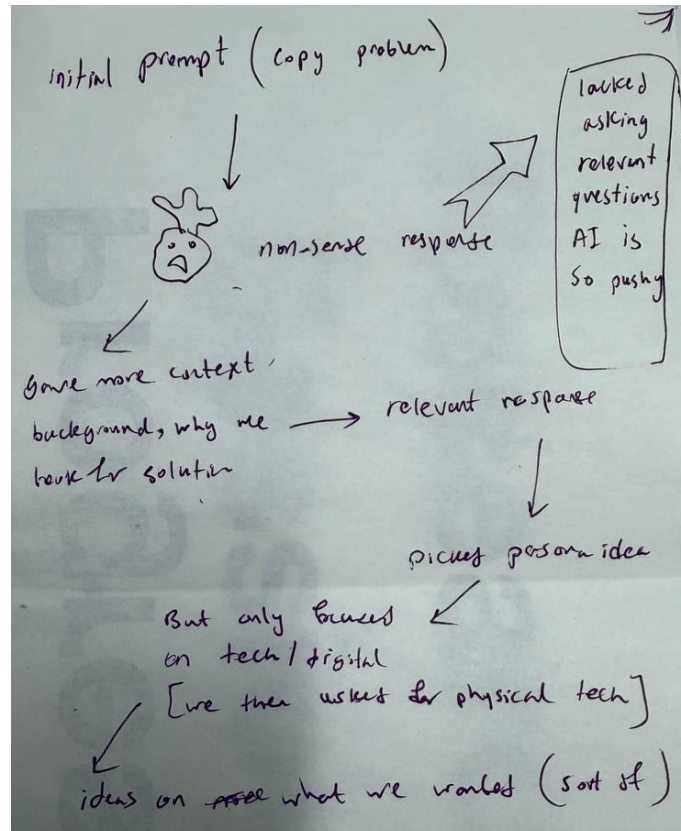


Figure A.9: Mind map: Participant 9

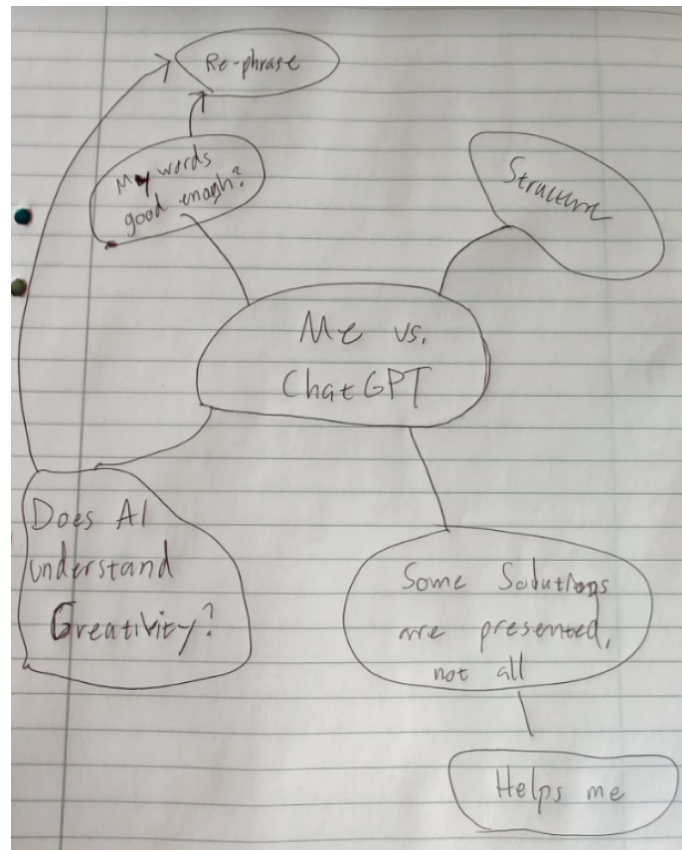


Figure A.10: Mind map: Participant 10

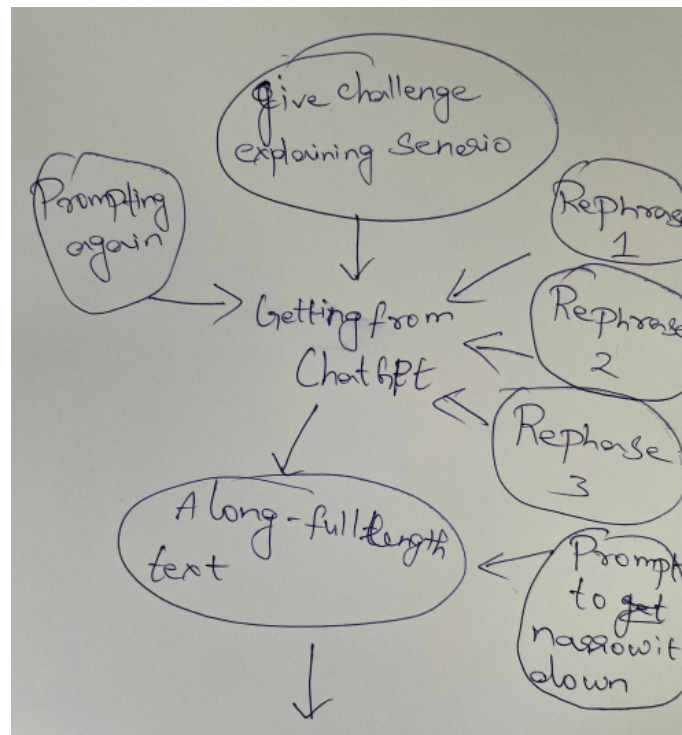


Figure A.11: Mind map: Participant 11

A.2 Love and break-up Letters

Participant 1 (Workshop 1)

Love Letter

Dear ChatGPT,
Working with you felt like having someone to bounce ideas off when things got stuck. You helped us move faster and gave us directions we probably wouldn't have thought of on our own.
Not everything you said was perfect, but that actually helped. It gave us something to react to and build from.
You didn't replace our thinking, but you definitely made it easier to keep going.
Thanks for that.
(written by chatgpt as I forgot how to write after using it)

Break-up Letter

Dear ChatGPT,
I think I've started to rely on you a bit too much. You're fast, always there, and full of ideas but sometimes it feels like I stop thinking for myself when I use you.
A lot of what you give sounds good at first, but then I realize it's a bit surface-level. I end up filtering more than I expected.
Still, it's strange. I can't really remember how I worked before you.
Maybe that's exactly the problem.

Participant 2 (Workshop 1)

Love Letter

Hi dear Chat, I love how you always have a response for me. Even the dumbest questions or the simplest tasks you are there for me and I truly appreciate you for that. The way you ask followup questions and elaborate is something that I adore. You are positive and try to keep me from doing wrong. The way you remember previous chats and projects always amaze me and keeps me going.

Break-up Letter

Dear chatgpt, the way that you keep writing the answers to my prompts are driving me crazy. You never stop talking?! Reasoning here and always trying to flatter me... Its driving me insane, and the way that you always pretend to know everything. Even the all mighty chat must at some point not know what to respond. "Sorry I dont know that" is that so incredibly hard to say?? Just try it once. And give it up with all those "- " jeez.

Participant 3 (Workshop 1)

Love Letter

Since the day I met you, my life has changed. I can't say I've become smarter, but you make me feel smarter, and that's one of the things I love most about you. Besides that, you help me understand other people and what they want, you make me see everything from different perspectives. Love u, baby.

Break-up Letter

Our relationship has become very toxic. I feel very dependent on you. I think I'm losing my essence and authenticity. I need some time away from you to think about which direction to take. Goodbye.

Participant 4 (Workshop 1)

Love Letter

Dear ChatGPT I have never thought about writing a letter to you so it is funny to talk to you in this way:)))
I appreciate you make me save time and give me several ideas in just a short time. Life has become easier with you .

Break-up Letter

Dear ChatGPT you've been a good partner for me but sometimes you just type me the thing I think about a problem just in different words and you do not come up with the ideas that are not biased with mines I think. Or in some cases you just give me sooo general ideas. and I need to talk you more and more and pass iterations. I just got lost in our chats sometimes I wish if you had bookmark feature maybe we could be partners yet!

Participant 5 (Workshop 1)

Love Letter

I appreciate how you have made my work easier and are helping me to make my workflow more efficient and time saving. I don't know how would I have handled things if you had not been invented.
Thank you,
Participant 2

Break-up Letter

It is really hard to say goodbye to you since you have been with me for around 2 years and have helped me a lot in my tasks and daily works from easy to hard ones.
I can't think about continuing without you and stop using you, even if we had some disagreements with eachother.

Participant 6 (Workshop 2)

Love Letter

Dear ChatGPT,

I love you for all the quick responses, Yu make my brainstorming sessions more varied and dont get tired of how much edits I want. You give my document backups so I can actually fake I did a research based on evidence. you make life easier

Break-up Letter

Please listen carefully to what I want to change, dont stick to one idea and presist on it when I need it to be changed. Say you dont know sth when you dont know and dont give me false of fake or made up results.

Participant 7 (Workshop 2)

Love Letter

Hi Mr ChatGPT,

I really love how creative you are, and how quickly you can think of something that I can't even imagine in such a short time; I love how patient you are, no matter how roughly I articulate my prompt, you seem to be able to catch the core idea and understand it very well.

Love you,
Participant 9

Break-up Letter

Mr ChatGPT,

I think our relationship will come to an end if you keep doing this! First of all, sometimes you make bullshit looks scientific and try to convince me to believe you. If I use my own brain to think for a moment, then I can figure out you are basically saying the same thing.

I hate that you are stuck with the creative loop, and always give me some ideas about certain features that are under the same theme. I tried to make you jump outside the box, you said yes, you will do it. But the next answer will still be the same theme.

Love you, Participant 9

Participant 8 (Workshop 2)

Love Letter

Hi! I like the way you collect information and the way you organise them. It feels like you know everything and so smart, no matter how tricky the question is, you can always find out an answer for it. You're the best tool for desk research.

Break-up Letter

I don't like you being so talkative and repetitive for sometimes, and sometimes your answer doesn't make much sense. And you're fluctuated, every time when I try to say something else, you will change your answer to adopt to me. It's annoying feels like you're just fitting in my preferences, not the true things.

Participant 9 (Workshop 2)

Love Letter

It does open my mindset to explore ideas I struggle to come up with, in way shorter time

Gives me confidence somehow, sometimes!

Break-up Letter

Too pushy - doesn't ask us for more context. It just pushes its idea thinking it's the best... so self-centered.

So text-heavy, list-based, and long scrolling

Participant 10 (Workshop 2)

Love Letter

Dear ChatGPT, my boo boo, I love how you take all my mental demand problems and make them go away in a satisfactory way. This design scenario was too much for my poor brain to handle at this moment and since you saved me so much demand I could direct more mental activity to my creative side and use that to come up with ideas based on the lovely things you said. I like that you structure this big scenario in a clear and precise way, it makes me want to kiss you even more since you just know me, thank you boo boo. And that you keep giving me more suggestions based on what I wrote, that is gold!

Break-up Letter

Listen ChatGPT, it's not working, I need to see other AI bots...I can not handle writing a big long text and then receive thousands of words that are not connected to what I wrote. Was it not clear that I wanted design solutions based on what I wrote? You started talking about something else, what are my words wrong? Too human for you? And also I do not like that you limit yourself to feeble emojis and lots of bullet points. I have to ask for you to make pictures and visual solutions? That should come with the answer you give, this is a betrayal! I deserve an AI how gives me everything at once. You are dumped, Chaty jerk.

Participant 11 (Workshop 2)

Love Letter

My dear ChatGPT,

It is very hard for me forget, how you easily solve my challenge in seconds which reduced my stress a bit. I love the way you asked my opinions and the suggestions you provided for me to choose which shows how committed you are with me. :)
In the end, you gave what I needed.

Break-up Letter

Hi again,

I see, that you give me lot a texts and lists which I am not conformatable at reading all the time. Reduce it as much as possible and ask my suggestions and opinions before you are doing a task not after you complete it.

B

Appendix 2

B.1 Introduction to Final Recommendations

The recommendations are structured into three different sections. Recommended knowledge, general recommendations and task-specific recommendations.

Section B.2 provides the rationale for the recommendations and is referred to in the recommendations for greater understanding. It was found in the workshops that designers' assumptions about AI often did not align with what these systems can actually do, making this information important to include. Therefore, it is recommended to read, or at least skim through this section, before proceeding to the recommendation sections to ensure a clear and accurate understanding.

The second and third sections include the recommendations. The second section, B.3, includes more recommendations that can be applied during all stages of the ideation phase. The third section, B.4, provides more detailed recommendations for specific stages of the concept exploration and ideation phase. The recommended knowledge (RK) includes three sections, highlighted and referred to in the recommendations as RK1, RK2, and RK3, for a more detailed understanding.

B.2 Recommended Knowledge (RK) about AI

During the workshops, a mismatch between designers' understanding of how AI works and how designers are affected by the tool was found. As a result, many of the recommendations are built on a foundation that explains AI's capabilities and limitations to utilise the tool to its full potential. Since research indicates it is more than just a neutral tool and can unconsciously negatively affect users, information about this was also deemed important to demonstrate and include in the recommendations. This section summarises the most important findings.

B.2.1 RK1: How AI mimics speech

As illustrated in Figure B.1, imagine you could learn a language fluently, but without understanding a single word. When someone would speak to you in this language, you would have no idea what they were saying, but by recognising words and knowing how sentences are typically structured, you would be able to answer correctly. You would also have no clue what you just said.

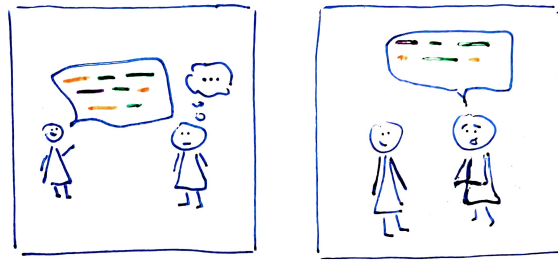


Figure B.1: Person speaking a new language fluently, but without knowing what is being said.

For a human, this is not possible, however in 1966, Joseph Weizenbaum set out to create something that could [91]. Through symbolism-based AI, he created the world's first chatbot, known as Eliza [92], illustrated in Figure B.2. It was a simple program that could translate text into a response based on the context and identified keywords.

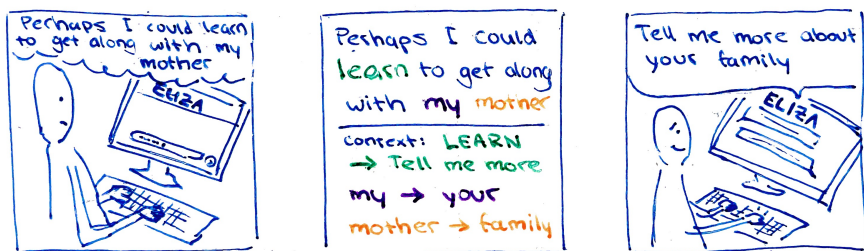


Figure B.2: Sketch of what how interaction with Eliza might have looked like. Based on one example from the report [91].

When discussing popular AI-chatbots today, LLMs are often what we think of, which are not based on symbolism, but instead on a language-driven approach, and are far more complex [93]. However, an LLM is also able to speak a language fluently, without knowing a single word. Additionally, what they also have in common is the type of attachment, trust, and faulty mental model that can form around the technology, which we will dive deeper into in the next section.

How does an LLM work?

An LLM is a large artificial neural network aiming to model how the brain works. However, it only scratches the surface of what the brain actually does and is capable of [94]. The neural network consists of parts of words called tokens and they are connected through edges with weights [45], [72]. At the beginning, the model's internal weights are random. During training, these weights are gradually adjusted as the model processes enormous amounts of text and learns which tokens are likely to appear in different contexts. A simplified illustration of reinforcement training is shown in Figure B.3.

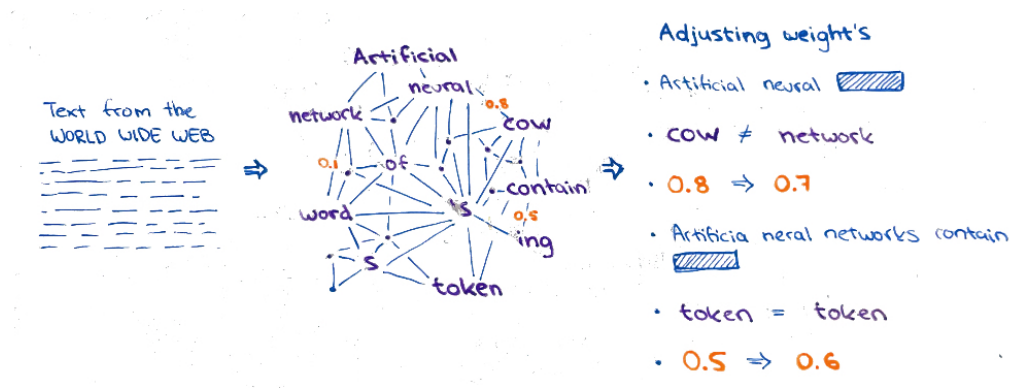


Figure B.3: A simplified sketch of how reinforcement learning changes the weights depending on whether the guess is correct or not. This results in a model that can structure sentences based on which word is most likely to appear next.

As a result, a model is created that can generate text that mimics how sentences are typically structured and adapt its content based on the keywords it receives in prompts. The first word is generated based on the prompt, and thereafter, what has been generated so far is also included to inform the next word. This way, the chatbot can respond in a way that fits the context and imitates the structure of our language.

Because the training is based on all types of data from the internet, the responses are not always appropriate and can contain all sorts of biases and stereotypes, as seen in Figure B.4. Therefore, Reinforcement learning through human feedback (RLHF) is required as a next step, where human preferences help guide the model toward responses that are more helpful, accurate, safer, or more appropriate [45], [72]. However, this process also increases certain biases and cultural values based on what is deemed appropriate. This is also a way to put safeguards in place and control what the chatbot replies and how. However, this can be broken through jailbreaking, which can be achieved by giving the chatbot personas and role-playing. [95].



Figure B.4: A sketch of how responses from an AI chatbot might change before and after adapting the model's weights based on what is appropriate to respond and not. This is done through Reinforcement Learning from Human Feedback (RLHF), a training method that uses human preferences to guide the model toward what it should respond.

What are hallucinations? Hallucinations are a commonly misunderstood term, according to the workshop results. The instances in which information was detected as incorrect or not specific enough were described as lies or as intentional, since the tool didn't want to admit its lack of knowledge. However, hallucinations are a feature of the system and appear when the probabilistic selection of certain words produces an incorrect or unsupported output [96]. Since the system generates outputs based on patterns, probability, and context, it cannot distinguish whether the output is true, false, or lacks substance [97]. Due to this, it has no ability to lie, and it cannot be open about its limitations and tell when it lacks knowledge, might be wrong or needs additional context. OpenAI supports the claim that their tool has no ability to tell truth from falsehood in their declaration of limitations [98].

Hallucination can also be confirmation that the requested instruction has been done, when it has been executed insufficiently or not at all. Additionally, the messages shown while waiting for a response are not hallucinations per se but could be seen as deceptive, since claims such as "thinking" or "comparing" do not accurately reflect what is being executed, as shown above. This could further imply to the user that a sense of intelligence exists, which is a design choice that creates a deceptive rather than a transparent UI.

Sycophantic design During the workshop, validation and complements were also thought of as hallucinations, but although they might be that too, this is mainly an effect of a sycophantic design. This is mainly introduced through RLHF [75], as seen in Figure 4, where responses change based on the preferences enforced by the data labellers. Sycophantic responses seem valid because the tone is confident, and validation is a human weakness that makes us turn a blind eye to irregularities and shortcomings in those responses. Furthermore, sycophancy drives hallucinations to happen, due to the system's design to give pleasing answers rather than truthful [87]. When looking for validation or confirmation, this suggests that the user hallucinates with the AI rather than just being

met with hallucinations from the AI [89].

B.2.2 RK2: AI Impacts on designers

After Weizenbaum tested his chatbot, Eliza, he stumbled upon unintended consequences when he revealed the program, and people refused to believe they had not been chatting with a human, or at least something intelligent [91]. As a result, they assigned much greater intelligence and trust to the program, despite its lack of human-like intelligence and ability to truly understand the user, a phenomenon later coined the Eliza effect. This is explained by neuroscientists as a mental model, an internal representation of how the things around us work [83]. This is necessary for us to understand, predict, and interact with the world. However, faulty mental models can have dangerous consequences. In Eliza's case, the perception of human-like intelligence led the user to form an exaggerated perception of the program's capabilities.

Although Eliza is completely different from LLM chatbots, a similar effect occurs. It is created by anthropomorphising AI, which means it is designed to resemble a human and give the user a perception of human-like characteristics [86], [99]. In LLMs it's built into responses, making them seem thoughtful, polite, and empathetic, which gives users the impression of being understood. Chinmulgund, Avanti, et al. found this trait to be the main contributor to the user anthropomorphising AI [85].

The choice to design the tool as a chat also gives it human characteristics, since it is a common way to communicate with other humans. The messages shown while loading responses, claiming it is "thinking" or "comparing", can further induce a false perception of what is happening behind the scenes and a sense of human-like capabilities. As with Eliza, this creates a faulty mental model and understanding of what the tool's capable of.

Impact of Perceiving the Tool to be Human

Anthropomorphism is used to make the tool more user-friendly and to heighten users' trust in it [85], [86]. This has been shown to reduce users' likelihood of questioning AI advice [86]. Sycophancy increases this risk by making responses feel valid, reliable, or contextually appropriate [75], [76], [87]. When a system is overly agreeable, affirming, or confidence-inducing, it can reduce the designer's critical distance from the output [76]. Additionally, studies show that this feature can create an echo chamber in which the user's ideas are amplified rather than critically questioned [77], [88], [89]. This creates a delusion, and in the worst case, psychosis. As the output contains bias, the types of answers received are also limited by the training data and by how RLHF has shaped the model. As illustrated in Figure B.5, this can limit the scope of ideas explored by the designer

while also having a feeling of shaping ideas and being in control. For these reasons, questioning output and being aware of how one is influenced can be challenging and may affect the development of ideas.



Figure B.5: A satirical sketch of how the user is in charge of making the final decisions, but only within the scope of what is being suggested.

AI output can further influence designers through cognitive bias and anchoring, which influence how one processes information and makes decisions [82], [83]. An early study on cognitive bias shows that even unrelated information can affect decision-making [84]. A popular example from this study illustrates how participants' guesses were influenced by the random number they obtained by spinning a wheel. Therefore, once AI is used for ideation, it may shape how the designer frames the problem, evaluates alternatives, or decides which ideas are worth developing. The designer should therefore be cautious about how and when AI is used, especially if it is applied early in the ideation process.

Risk of cognitive decline: Creativity and critical thinking

A faulty understanding of capabilities, sycophancy and heightened trust can also increase the designer's reliance on the tool [76], [77]. It may be tempting to use AI to reduce mental demand, but not being challenged can lead to cognitive decline and the loss of important skills, such as critical thinking, judgement and decision-making [78], [79], [80], which are essential to possess as a designer. Additionally, creativity is a skill [81], that we risk losing if it is not exercised. Designers should therefore be wary of which responsibilities and forms of work they delegate, since

giving away too much of the creative process may reduce both creative practice and control.

When using AI to improve or refine ideas, the introduction of biases and generic solutions also risks flattening the ideas, stripping them of their uniqueness. McMeel et al. argue that, because AI cannot bend or break rules, it lacks the potential to be truly creative [43]. They also point out the value of mistakes and human abilities to leverage errors for creative gain.

B.2.3 RK3 Output validity and Efficiency of Context

How context is used

LLMs use transformers, which were developed for Google search, a model that predicts which words in a sentence are most important rather than focusing on every word in the sentence [71]. This way, context can guide the model towards more relevant and specified responses [45], [72]. The need for detailed context comes from the many possible meanings that words can take when combined. For example, a search for “design solutions” could refer to architecture, games, furniture, services, or many other domains. By providing more precise information, the AI model is better able to retrieve or generate data that matches the intended request. However, more context does not eliminate hallucinations. These are part of how AI systems operate and can arise from the way information is associated and generated within the model. There are also limits to how much context can improve an answer.

How efficient context is and how valid responses are

Research finds that AI advice often fails to adapt to context and situation, yet creates the impression of being insightful. Dergaa et al. found that using AI for medical advice resulted in inaccurate assessments, yielding generic, similar responses that failed to adapt to the different contexts and scenarios [73]. Another study from Harvard found that context had little effect on the validity of business advice [74]. Prompting different LLMs with different scenarios and amounts of context showed that similar advice was given regardless. They found that the main factor driving variation in advice was the order in which options were presented, choosing one of the top options regardless of the situation. Although the studies do not assess the validity of design solutions and ideas, they assess the ability to evaluate situations and provide advice accordingly across two distinct domains. This highlights the model’s limitations in analysing, resonating, and adapting suggestions, which may have negative impacts in ideation as well.

In addition to the response’s accuracy, context was found to make advice look more tailored and specific, even though it wasn’t [74]. Dergaa et

al. noted its capability to extend empathy towards the patient [73]. As explored further in the section above, this is likely due to the sycophantic design of responses, which tricks us into thinking something is insightful when it is delivered in a personal, authoritative, and confident way [75]. Anthropomorphism has also been shown to heighten users' trust in AI advice, which reduces critical thinking [86]. As a result, the content may appear adapted and, at times, even be valid. However, the relevance and quality of its suggestions and advice cannot be reliably trusted, and the designer should therefore always reflect upon the output and stay critical of it.

Instead, one can use AI to ask counter-questions about the design, prompting the designer to reflect and improve it using their own knowledge and judgement. However, be aware that the responses, even when challenging, are still tailored to please the user. In the end, the tool merely imitates how advice, reasoning, and empathy are expressed and cannot adapt advice to context due to the absence of true understanding.

B.3 General AI recommendations across the Ideation Phase

B.3.1 Recommendation 1: Direct the output through context

A clear context and goal of the output are required to guide the response in the right direction. By giving more precise information, the AI model can find data that matches your request better. This does not eliminate hallucinations, which is part of the system and exist due to how data is connected in the AI model, explained in RK1. Additionally, a new studies have shown that context has a low effect on how valid and specific output is, which can be read about in RK3. What it mainly affects is an increased feeling of the response being tailored to the users request. This is due to the sycophantic design, which tricks us into feeling something is insightful when it is being delivered as personal and with confidence.

Four key prompt components and how they affect the output

As described above, context influences which words in the artificial neural network are most likely to be considered relevant. This knowledge can be used intentionally when deciding what to include in a prompt. Based on the workshop and recommendations from Nielsen Norman Group [70], four components are especially important:

- **R1.1 - Context:** Describe the situation and what you are looking for. This can include who you are, what you are working on, and what the end goal is. For example, if “UX designer” is included in the context, the output is more likely to focus on words, concepts,

and solutions associated with UX design. This is explained in more detail in RK3.

- **R1.2 - Instruction:** Explain what the AI should do. This defines the task, such as generating ideas, summarising insights, challenging assumptions, or creating alternatives. As RK1 shows, AI cannot truly understand or faithfully adhere to instructions. Therefore, you cannot trust that all instructions are fulfilled.
- **R1.3 - Constraints:** Provide rules or boundaries for the output. This can include what to avoid, how detailed the answer should be, what format to use, or which perspective the AI should take. As mentioned in R1.2, there is no guarantee that constraints are followed.
- **R1.4 - Wanted output:** Explain what your desired structure, tone, or type of output is, by providing an example of what you want the response to look like.

Be aware: Research suggests that adding context has a relatively low effect on the objective quality of AI output, and is explained in RK3. What it primarily affects is an increased feeling that the response is tailored and specific, due to anthropomorphism and sycophantic design. Context can therefore be used to direct the response, but should not be mistaken for an insightful or high-quality one.

B.3.2 Recommendation 2: Challenge and stay critical of AI responses

One drawback of using AI is that it cannot be given responsibility for mishaps or decisions you have made based on its output. At the same time, as described in RK2, the format of the response can make you attribute more validity to the output than it actually deserves. This can make it difficult to remain critical of AI-generated output and maintain control of the design process. The usage can also easily slip away from being controlled to being more trusting than was first intended. Therefore, it is important, as a designer, to challenge outputs to make informed decisions and be aware of how they affect both oneself and the design.

- **R2.1 - Enforce critical thinking with knowledge:** Knowing that the system is not capable of true reasoning and its creative limitations can help you stay critical of the output and challenge what you are being given. However, there is a crucial difference between knowing and understanding. We can sometimes make decisions that go against our better judgment, even when we know better. Understanding, however, strengthens our ability to act against impulse, habit, or external pressure.[cite!!]. To gain a greater understanding of how AI works, it is recommended to at least skim through RK1. In addition, to understand the effects

that AI can have on the designer, it is recommended to at least skim through RK2.

- **R2.2 - Dehumanise responses:** Lowering the confidence, validation, and human format of AI output can make it easier to use as a tool, avoid attributing intelligence that doesn't exist to responses, and therefore avoid giving it unwanted authority. RK2 explains how the current design of the tool tricks the brain into perceiving greater abilities than it has, and how this affects you and the design.

Be aware: Designers who do not actively counteract sycophantic behaviour risk entering a feedback loop where AI validates whatever they propose, and they in turn accept whatever AI suggests. The result is a narrowed, mutually reinforcing idea space rather than genuine creative exploration.

B.3.3 Recommendation 3: Change the settings to customise responses and memory

Settings can be used to customise the tone and structure of responses, repetitive context, and memory management. For UX designers, these settings can be used to get responses that are more relevant and easier to digest. The suggested settings are specifically to make the design process easier.

It is important to note that these customisations can make the collaboration easier, but they do not eliminate hallucinations or sycophantic responses. What this is and how it works is explained in more detail in RK1.

- **R3.1 - Response Structure:** The point of using AI in design is to make the process more efficient. Therefore, as a designer, you do not want to be overwhelmed by word gibberish and having to sort through long answers, a challenge raised during the workshop. To facilitate this, settings can be used to make responses more concise, using fewer lists and emojis. These settings make the response more to the point and less repetitive. However, repetition can still be found within multiple responses over time.
- **R3.2 - Response Tone:** Tone can be adjusted to be less validating and complementary, and more direct and critical. During the workshop, many designers raised the issue that validation was problematic, especially when it made a bad idea seem good. This feature stems from the sycophantic design of AI chatbots, and it can be reduced by adjusting the tone. However, this does not change the core of this feature, and it will still produce text to your satisfaction and in a way that seems specific and convincing. Sycophantic design is explained in RK1, and more about how it affects the designer in RK2.

- **R3.3 - Memory Management:** There are two types of memory: shared memory and stored memory. Knowing how to use this can help you get more specific responses without losing access to information that is appropriate to share between chats.
 - Shared memory: If this feature is turned on, the AI has access to separate chats and uses those as context. Be aware that information from separate chats can be used without your knowledge or instruction to do so. Since you can't know how or when other chats are being used for context, it can influence your chats in unwanted ways. For example, specifications for different projects can be mixed.
 - Memory storage: Context and customisation can also be added as “memories” to a separate memory storage that all chats have access to, regardless of whether shared memory is turned on or off. Memories can be added or deleted in the chat by sending a prompt containing “Write to memory: ...”. The memories can be viewed and deleted in settings, but not added to or changed.

Be aware: These customisations do not eliminate sycophantic design or hallucination. They reduce their surface expression, which lowers unearned trust, but the underlying model behaviour remains. This can cause responses to still confirm the designer's view, but less explicitly, making it more difficult to detect.

B.4 AI recommendations for specific stages of the ideation phase

This recommendation was created to meet the designers' needs while also accounting for potential negative impacts and risks. As a result, three different ways that designers can use AI to get started, as a warm-up or inspiration, were created, while also avoiding high-fidelity ideas and solutions early on.

Through the workshop, we discovered that many liked to get started using AI, either as a warm-up to get into the mood or to get inspiration. However, when using AI chatbots for this, it was also observed that responses tend to go straight to ideas and solutions. This should be avoided, per the design principle of avoiding high-fidelity sketches and mock-ups early on, since they are more difficult to improve. Before AI, this was mainly due to the time required to make high-fidelity prototypes. However, finished-looking designs are also more difficult to reflect upon and to detect weaknesses in. As a result, this influences the ability to reiterate and make valuable changes to the design.

As explained in RK2, AI can also negatively affect designers by introducing biased and generic ideas. Whether you are aware of it or not, this will

influence later decisions because it registers as a cognitive bias. Therefore, using AI at the beginning of the phases can have greater impacts, and it should preferably be avoided.

B.4.1 Recommendation 4: Use AI to Spark Inspiration While Keeping Creative Control

- **R4.1 - Use AI to create Mood Boards:** Ask it to generate a mood board of words, images, or concepts related to the problem. Use these as triggers to spark your own associations before moving toward ideas. However, be aware that the output holds biased cultural values and stereotypes, while also tending to be generic. As explained in RK2, what you're being exposed to will influence your cognitive bias and risk of being brought into your final solutions.
- **R4.2 - Force the output to create low-fidelity sketches:** Provide instructions to generate low-fidelity sketches rather than images. This tackles the issue of high-fidelity, finished-looking ideas, which limits the designer's ability to be creative and iterate on ideas. Since all ideas will influence cognitive bias (RK2), lower-fidelity ideas may help the designer balance it against their own input.
- **R4.3 - Use AI as a Starting Point, but Leave Gaps for Imagination and Problem-Solving skills:** Ask for ideas of how to solve a problem, but instruct the output to blank out descriptive words and fill in the gaps using your own imagination and problem-solving skills. This helps you, as a designer, get in the mood for problem-solving while being less inclined to go in a certain direction. It also forces the designer to stay in control and contribute with their own creativity.

Be aware: The sycophantic design of AI responses reduces the designer's ability to be critical, making outputs easier to accept and harder to question. AI bias, therefore, poses a greater risk of influencing the final design more than it might initially appear. Importantly, designers are affected by cognitive biases even when AI responses are not in the form of direct ideas or solutions.

Ideation and Exploration The following recommendations cover how to effectively use AI to generate ideas and explore them. However, as described in AI as a starting point and RK2, it is not recommended to begin ideation with AI due to biased, unoriginal output, its potential to introduce cognitive bias, and its tendency to narrow the designer's creativity. Therefore, it is better to start by forming your own ideas and solutions, and then use AI afterwards to get additional ideas and alternatives.

Given this, the following recommendations offer tips for generating diverse ideas and thinking outside the box. Following these recommendations

becomes even more crucial if the designer's own ideas have not been explored beforehand.

Be aware: Avoid giving AI responsibility for both generation and selection. If AI creates the ideas, the designer should choose. If the designer generates the ideas, AI can help develop them. Delegating both to AI removes the designer's creative contribution from the process.

B.4.2 Recommendation 5 - Generate Diverse and Rough AI Ideas, Then Explore Independently

- **R5.1 - Ask for a large number of diverse ideas:** Start by asking for a large and diverse set of ideas and specify how many, for example, 20. Explicitly ask for ideas across completely different approaches. Quantity and variety first, quality second. Although there may still be repetitions present, the quantity forces the response to find new variations. The quantity also allows for more “rough” ideas to be discovered, displaying the most probable, biased and generic ones at the top.

- **R5.2 - Start from the bottom: Explore the Least Promising Ideas First:** Before evaluating any of the ideas, start reading from the bottom. If there is a time constraint or the quantity is found overwhelming, select the last 5 ideas and do not explore any of the top ideas.

AI-generated ideas tend to get less applicable the further down you go, making the top ideas stand out as the better ones. Starting at the bottom may therefore seem contradictory to the goal, but as explained in RK3, this helps the designer stay in control. When working with ideas of lower fidelity, the designer is better equipped to react, combine, iterate and make changes to the generated ideas.

- **R5.3 - Work on ideas outside the chat, preferably in a canvas space:** During the workshop, designers raised the issue that it was difficult to get an overview and combine ideas. Therefore, we suggest copying ideas to a FigJam or Miro board to get a better overview. You can also use an AI Miro board if you want AI and canvas in one place, and to build on ideas using AI in a canvas space.

Be aware: Avoid giving AI responsibility for both generation and selection. If AI creates the ideas, the designer should choose. If the designer generates the ideas, AI can help develop them by questioning the designer and their design choices, as explained in the next recommendation. Delegating both generation and selection to AI removes the designer's creative contribution from the process.

Converging and refining ideas with AI

When converging and refining ideas, it is important to remember that the output, and therefore the suggested changes, are biased. You might, therefore, accidentally make your solution more generic and lose the uniqueness when using AI to refine it. Additionally, AI should not be used for advice, since it does not have the capability to understand requirements or reason toward a better solution. This is explained in RK3.

B.4.3 Recommendation 6 Do not Use AI for Advice, Use It to Challenge Your Ideas

- **R6.1 - Do not use AI to narrow down the selection or select the best idea:** Due to the system's inability to reason, it is not a good idea to ask for advice on which idea to pick or to narrow down the selection. According to a new Harvard study, AI advice is driven by the order of suggestions and tends to prioritise the top ideas regardless of validity. This means the advice is not based on logic but rather on the order in which options are presented and sometimes on how common or popular they are. As a result, selections may also be narrowed to specific themes based on the order, even if not prompted to do so. The same goes for asking AI which idea is best. Additionally, suggestions and refinements are biased, which risks being introduced into the ideas and selections, making them more generic.

If asking for advice, despite this recommendation, there is a high risk that responses will seem valid, and this recommendation will be overlooked. However, as explained above and further in RK2, the deceptive and convincing nature of responses reduces our ability to assess their validity. Therefore, the easiest way to avoid being manipulated is not to use it for this.

- **R6.2 - When refining ideas with AI, ask it to challenge you:** It is not recommended to use AI for advice or decision-making. However, it can instead be used to challenge your ideas and, in this way, help you question design choices, identify possible improvements, and clarify which directions you think are worth exploring further. To do this, break your communication with AI into steps. At each step, ask it to challenge what you have rather than build on it. At the same time, ask AI to be critical and, as described in RK2, do not accept AI outputs without questioning them either. Push back, probe, and stress test in both directions. This critical and challenging stance can also be changed in settings or stored in memory, as described in R3.
- **R6.3 - Begin with a new chat to start fresh:** When exploring specific ideas, make sure that shared memory is turned off in the settings and open a dedicated new chat for each idea. A fresh

chat gives each direction a fair and focused exploration without interference from earlier ideas.