

Design Master Thesis

DESIGNING UI WITH AI-GENERATED VISUAL HINTS

MSc Digital & Interaction Design

Virginia Maria Gabriela Arnaboldi

994900

Politecnico di Milano

Supervisor **Davide Spallazzo**

A.Y.2023-2024

Design Master Thesis

DESIGNING UI WITH AI-GENERATED VISUAL HINTS

MSc Digital & Interaction Design

Virginia Maria Gabriela Arnaboldi

994900

Politecnico di Milano

Supervisor **Davide Spallazzo**

A.Y.2023-2024

ACKNOWLEDGMENT

Eng.

Those years at the Politecnico were long and hard. It wasn't until I left Milano and applied my skills in a professional environment that I fully appreciated the education I had received. Looking back, the decision to specialize in D.I.D. was decisive, broadening my perspective on UX design's importance, cultural interactions, and the life experiences that deepen our understanding of people (who are not mere users of the interfaces we design).

I would like to thank **Prof. Spallazzo**, whose advice at the end of my undergraduate studies reshaped the academic path I had envisioned and also guided the end of my Master's journey.

I would like to thank **my father** for the suggestions that still guide my choices and for supporting my decision to pursue this course of study from the beginning. **My brother and sister** deserve special thanks for their constant support, encouragement and the long phone calls over the years. Thanks also to **Dario** for his job-hunting advices and to **Evangelia** for the shared moments of relaxation.

Of course, I cannot forget **the people who have been with me on this journey**. In a short time, we have created a bond that withstands the miles that separate us, built through shared tears, late-night work sessions, and existential reflections that we still share.

Lastly, my appreciation goes to **those in a foreign land who believed in me**, who continue to seek and listen to my opinions, who have given me opportunities to broaden my horizons, and who have allowed me to demonstrate my abilities beyond the academic world.

My sincerest thanks to all.



Ita.

Questi anni al Politecnico sono stati lunghi e difficili. Credo di aver veramente saputo apprezzare le conoscenze acquisite solo dopo aver iniziato a metterle in pratica nel mondo professionale, allontanandomi da Milano. Ma, tramite un'analisi retrospettiva, scegliere D.I.D. ha modificato i miei orizzonti, capendo l'importanza dello UX design, ma anche dello scambio culturale e del fare esperienze di vita che aiutino a capire le persone (che non sono solo utenti delle interfacce che progettiamo).

Voglio ringraziare il **Prof. Spallazzo** il cui incontro alla fine del triennio ha completamente modificato quello che pensavo sarebbe stata la mia futura carriera accademica e che ha accompagnato anche il termine del mio Master.

Voglio ringraziare **mio padre** per i suggerimenti che ancora guidano le mie scelte e per aver supportato fin dall'inizio la decisione di questo percorso di studio. E soprattutto **i miei fratelli** per il costante appoggio, supporto e le lunghe telefonate in questi ultimi anni. Un grazie anche a **Dario**, per i suggerimenti per trovare un lavoro, e ad **Evangelia**, per le vacanze e i momenti di relax offerti e condivisi.

Naturalmente non posso che ringraziare **le persone che mi hanno accompagnato nel percorso**. In breve tempo abbiamo creato un legame che resiste ai chilometri di distanza che ci separano, costruito tramite le lacrime, le notti di lavoro e i dubbi esistenziali che abbiamo condiviso e tutt'ora condividiamo.

Infine, voglio ringraziare **le persone che in un paese straniero hanno creduto in me**, che tutt'ora chiedono e seguono il mio parere, che mi danno la possibilità di ampliare la mia esperienza e che hanno permesso di dimostrare a me stessa il mio valore fuori dall'ambito accademico.

Un sincero grazie a tutti voi.

TABLE OF CONTENTS

| | | |
|------------|---|-----------|
| 0.0 | Abstract | 8 |
| | Eng. | 8 |
| | Ita. | 8 |
| 1.0 | Introduction | 10 |
| 1.1 | Topic selection | 12 |
| | Generative AI | 12 |
| | Text-to-image generation | 14 |
| | Target audience (UI Designer) | 15 |
| 1.2 | Methodology | 16 |
| | Literature: selection criteria | 18 |
| | User involvement: method and objectives | 19 |
| | Digital ethnography: methodological limitations | 21 |
| 2.0 | State of the art | 26 |
| 2.1 | Discussion | 28 |
| | Literature conclusions | 28 |
| | Survey results | 44 |
| | Ethnographic research findings | 54 |
| 2.2 | Triangulation results | 60 |
| | Define the role in the design process | 60 |
| | Facilitate interaction with the tool | 62 |
| | Propose new evaluation criteria | 63 |

| | | |
|------------|---------------------------|-----------|
| 3.0 | <i>Conclusions</i> | 64 |
| 3.1 | Process | 66 |
| 3.2 | Results | 67 |
| | Critical analysis | 67 |
| | Limitations | 68 |
| | Future steps | 69 |
| 4.0 | <i>Appendix</i> | 70 |
| 4.1 | List of reference | 72 |
| 4.2 | List of figures | 78 |
| | List of figures | 78 |
| | List of tables | 79 |
| 4.3 | Appendix | 80 |
| | Appendix 01 | 80 |
| | Appendix 02 | 81 |

ABSTRACT

Eng.

In the context of rapid technological advancement, the field of design cannot ignore the significant advantages offered by AI, especially in terms of data accessibility, given the cognitive and temporal limitations of the human mind. Recent applications in visual arts and design have highlighted the lack of guidelines and relevant metrics tailored to these fields and professionals' needs. This thesis examines the potential **value and contribution of AI-generated visual cues to the UI design process** and the **evolving role of the designer**.

Both strengths and limitations of AI technology suggest that its generations are not yet suitable as final products in UI design. Instead, these can be a **source of inspiration in the early stages of ideation**. Investigating interactions with AI models and prompt construction is essential not only to effectively communicate the designer's intent but also to understand the added value of originality and diversity in the results. The unpredictable and vague outcomes can enhance creativity, and combined with each designer's subjective sensibility, can mitigate the risk of copyright issues. The overall aesthetics of the visual hints influence the design process more than their realism. Suggested **metrics based on user needs** can lead to effective qualitative assessments and deepen our understanding of the limitations in generating the specific theme.

The research employed a multi-method approach. It combined theoretical knowledge and assumptions from academic studies with surveyed opinions and experiences from professionals. The study validated current applications, tools' capabilities, and patterns of ideas in real-world professional settings through digital ethnography on the YouTube platform, which promotes practice and information sharing. It has been assessed that the designer's **informed evaluation in an iterative and non-linear process remains essential** for creating user-centered designs and cannot yet be replicated. Nevertheless, AI technologies have the potential to **change collaboration dynamics and the goal of solution development**. The impact will be determined by the intended applications and the ability of designers to adapt their skills.

Ita.

Nel contesto di rapido sviluppo tecnologico, l'ambito del design non può ignorare i vantaggi offerti dall'IA in termini di accessibilità di dati, soggetti nella mente umana a limiti cognitivi e temporali. Le recenti applicazioni in campo di arti visive e design sottolineano l'assenza di linee guida e metri di giudizio rilevanti per il campo specifico e i bisogni del professionista. Questa tesi, perciò, esamina il potenziale **valore e contributo degli spunti visivi generati al processo progettuale di interfacce digitali** e, di conseguenza, il **dinamico ruolo del designer**.

Entrambi punti di forza e limitazioni della tecnologia non supportano l'applicazione delle generazioni nello UI design come prodotti finiti. Questi possono essere **fonti di ispirazione nell'iniziale fase di ideazione**. Indagare l'interazione con i modelli di intelligenza artificiale e la costruzione del prompt è essenziale non solo per una efficace comunicazione dell'intento del designer, ma anche per comprendere il valore aggiunto dall'originalità e dalla varietà dei risultati. I risultati vaghi e inaspettati possono supportare il processo creativo e, con la profonda soggettività che distingue la percezione di ciascuno, dovrebbero mitigare il rischio di plagio. Il valore estetico degli spunti visivi influenza più del realismo del suggerimento. La definizione di **criteri valutativi a partire dai bisogni dell'utente** può portare ad una efficace valutazione qualitativa standardizzata dei risultati ottenuti e approfondire la nostra comprensione delle limitazioni nella generazione dello specifico tema.

La ricerca è stata condotta utilizzando un approccio multi-modale. I presupposti sulla quale si basano gli studi accademici e la conoscenza teoretica prodotta sono stati raccolti e confrontati all'esperienza professionale e alle opinioni indagate tramite un questionario. È seguita una validazione delle correnti applicazioni, dell'effettive capacità degli strumenti e modelli di percezione all'interno del reale contesto professionale attraverso digital ethnography tramite la piattaforma di YouTube, che promuove la condivisione di informazioni e pratiche. La **valutazione del designer in un processo progettuale iterativo e non-lineare rimane essenziale** per creare user-centered design e, al momento attuale, non può essere replicata. Ma la tecnologia **trasformerà i ruoli nella collaborazione e l'oggetto della soluzione sviluppata**. L'impatto dipenderà dalle sue modalità di applicazione e dalla capacità dei designer di adattare le loro competenze al nuovo contesto.

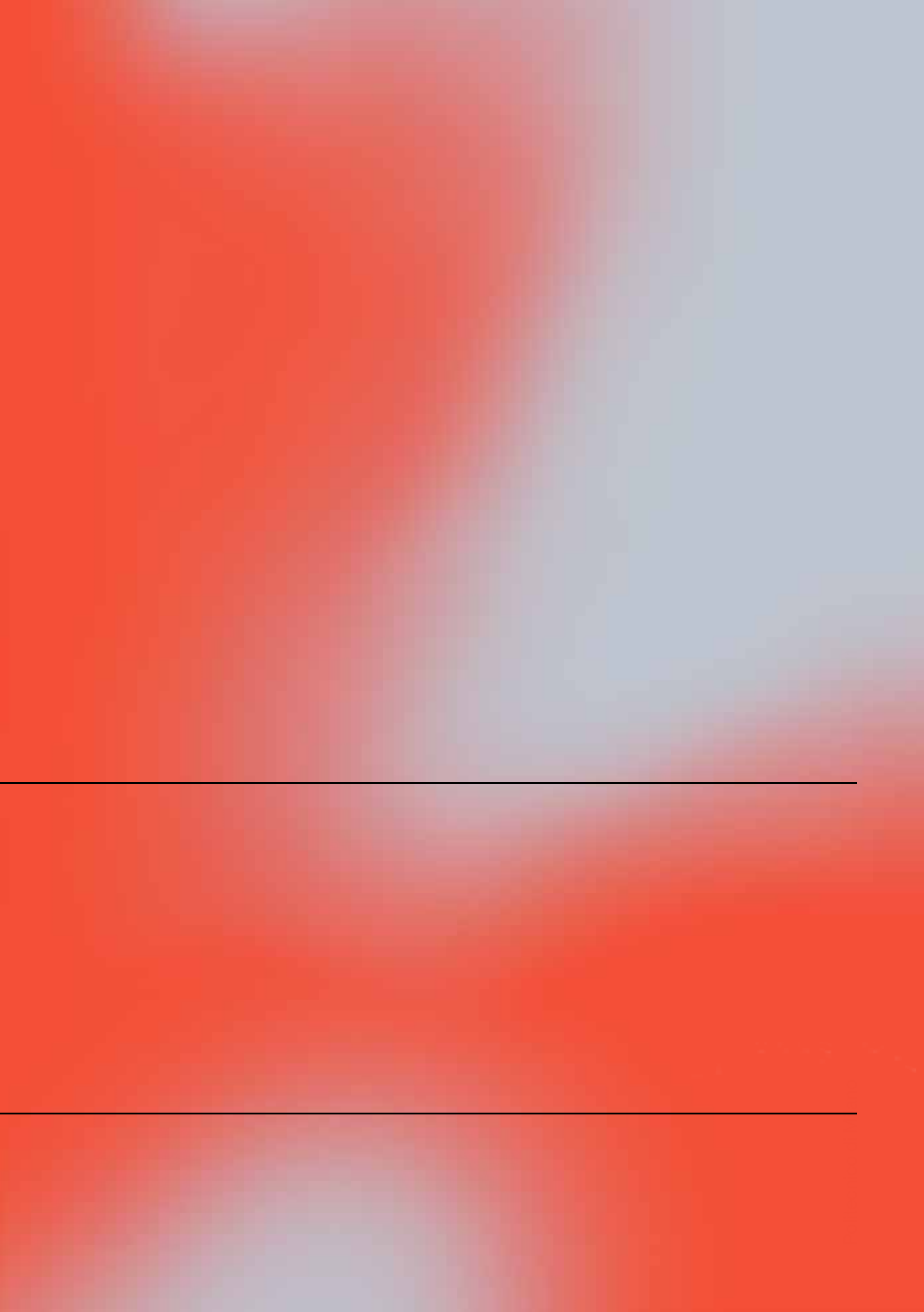
INTRODUCTION

1.1 Topic selection

Generative AI
Text-to-image generation
Target audience (UI Designer)

1.2 Methodology

Literature: selection criteria
User involvement: method and objectives
Digital ethnography: methodological limitations



Topic selection

Generative AI

AI's influence has extended across diverse fields, suggesting that in the future it's unlikely that any area will remain untouched by its impact. Those who do not embrace it risk being left behind. As digital technology rapidly advances, designers are challenged with continuously improving and enhancing our interactions with the new tools. The design sector is also being asked to evolve and integrate this growing technology into its practices. It is therefore crucial to explore how the adoption of AI can benefit and drive the sector forward.

This thesis does not have the final goal to address the definitions of “intelligence”, “culture” and “creativity”, nor does it explore whether these traits, traditionally seen as characteristic of humans, can be attributed to machines. Such discussions, including the relationship between art, creativity, innovation and the role of AI as a catalyst and driver for economic growth, are beyond the research scope of this work and will not be pursued (Urietti, 2023; Manovich, 2022). Instead, this study takes inspiration from Manovich's perspective (2022) in the fourth chapter of “Artificial Aesthetics: Generative AI, Art, and Visual Media” suggesting a shift in focus from questioning AI's creative capabilities to exploring its potential contributions across various artistic disciplines, including art, architecture and design. The re-orientation forms the basis of this exploration.

The enormous advantages offered by Artificial Intelligence in making data accessible cannot be denied, as it overcomes the cognitive and temporal limitations imposed by the human mind. Its ability to quickly consume and analyse vast amounts of data could also be an advantage in the field of visual design, where experts can spend long periods of time researching for inspiration, sketching out sets of ideas and defining solutions (Cai et al., 2023; Hanafy, 2023). AI's potential to aggregate extensive online information could provide a solution to the time constraints faced by designers, acting as a valuable resource for insights into emerging trends and serving as a comprehensive reference tool (Mozaffari et al., 2022).

Text-to-image generation

First emerging models of interaction for AI image generation were based on textual input. This approach democratises access to creative expression by eliminating the need for specialised artistic, pictorial or technical skills (in areas such as photography, digital editing, 3D modelling and rendering). The choice could be based on the assumption that the ability to articulate our thoughts and visualise something in our minds based on a simple description is one of the most innate human activities, and therefore the belief that it will facilitate our interaction with machines.

<<Language mediates as a conduit between how we think in our mind and how we externalise it.>> (Karahan, Aktaş and Bingöl, 2023)

But the complex link between the names we assign to objects and the objects themselves is a human convention. It's commonly presumed that words and images share a direct causation, or at least correlation, yet their association is far from straightforward (Koh, 2023). It's therefore challenging to replicate this human ability in a system, generating images that reflect the essence of a description (Frolov et al., 2021). Rather than thinking in terms of pieces of information, computer vision techniques focus on the most fundamental element of an image: the pixel (Miranda, 2021). The construction of an image can be conceptualised, from the point of view of the model, as the organisation of these basic units, chosen for their potential collective resemblance to identifiable objects. Advanced models, such as large transformers, are capable of generating novel pixel arrangements that offer variations on the original subject, using contextual cues to form recognisable shapes (Chen et al., 2020). As a result, the subject of the images remains consistent across representations, yet each image is unique, reflecting the probabilistic nature of generative models (Siddiqui, 2023). Training pixel sequences leads to the creation of coherent images and samples. However, simply recognising patterns or repeated elements isn't enough, because pixel sequences don't have explicit labels that identify the subject they represent (Chen et al., 2020). A crucial step is labelling, which associates pixel configurations with the names of the objects they are supposed to portray, essentially enabling the system to recognise them in the same way as a human observer might.

The term “**text-to-image generation**” describes computer-based methods that turn human-written prompts into visual representations with the same semantic meaning (Miranda, 2021). In essence, they aspire to materialise any concept that can be envisioned within our minds and then described. They can deliver a wide range of visual content, including realistic images (Hanafy, 2023). The progress in terms of visual realism, diversity, and semantic alignment that AI models have seen in recent times is significant, but the field still faces several challenges (Frolov et al., 2021). Given its inherently abstract and undefined nature, the same textual description can evoke different visual interpretations, even in each individual (Karahan, Aktaş and Bingöl, 2023). The process of translation inherently lacks specificity: a single caption can correspond to countless plausible images (OpenAI, 2021). Therefore, articulating something specific and accurately capturing the image somebody has in mind is a significant challenge. The difficulty of communicating effectively with the tool has been highlighted as a critical issue by both researchers and users. To facilitate interaction with generative tools, features such as auto-completion, repositories of previously used prompts, and dashboard interfaces for selecting presets have been introduced. However, **these aids can**

unintentionally homogenise practices and limit the exploration of new possibilities, undermining the flexibility and broad potential of text-to-image generation (Sanchez, 2023). On the contrary, this inherent ambiguity opens the door to misinterpretation and unforeseen results. Sometimes, the mere adoption of a new point of view can inspire new ideas, leading to innovative solutions and creative visuals.

It's important to note that this research does not imply the author's endorsement of text-based interactions as the optimal user experience (UX) method, nor does it predict this as a future trajectory. This dialogue parallels the challenge of constructing a questionnaire, where the choice is between open-ended questions and pre-determined options in closed-ended ones. The latter are more straightforward for respondents, but offer limited insight into their nuanced perspectives and thought processes. Consequently, this research focuses on this fundamental form of interaction, as provided by current tools, with the aim of **exploring both its challenges and its distinct advantages**.

Target audience (UI Designer)

While exploring the use of generative AI, a noticeable research gap has emerged regarding its **integration into UI design**, despite various architectural and product design experiments. Success in the latter areas has been demonstrated by the ability of the tools to generate compelling and ultra-realistic visualisations. Initial investigations into the design of digital interfaces have been made and are detailed in the following sections.

This research was inspired by a Medium article that proposed a different use of AI-generated images as creative inspiration for the development of a mobile application (Babich, 2023). Other examples in UI design have been documented online, in particular using the prominent tool Midjourney. The aim is therefore to explore the unique contributions that AI-generated images can offer across the design process, **beyond simply producing visually stunning results for immediate use in graphic layouts**. Recent applications in the visual arts and design domains underline the urgent need for new, relevant evaluation criteria tailored to these fields. The definition of such criteria, based on the needs of designers, could facilitate a standardised, qualitative assessment of results (Lee et al., 2023).

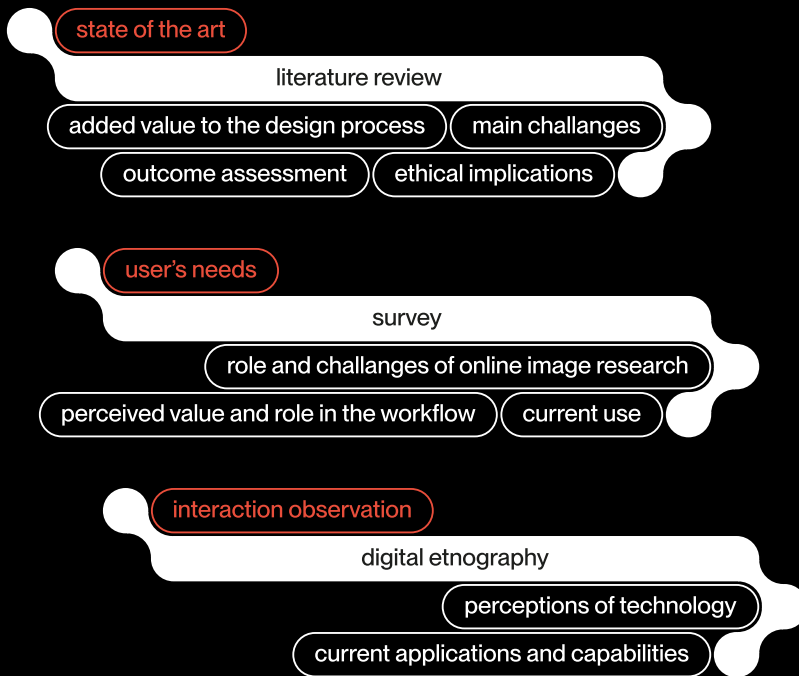
The potential role of this technology, the interaction with tools (including prompt construction) and criteria for model evaluation were explored. A literature review was conducted to map the research that has been done in this area, although it goes beyond traditionally validated sources. Designers were engaged through a survey designed to verify or challenge findings in the existing literature and to collect complementary data. At the same time, the digital ethnography method gathers insights into users' perceptions through content and opinions, particularly from YouTube, which drives knowledge sharing on the web.

The relevance of this research lies in its ability to highlight knowledge derived from both academic perspectives and non-scientific sources. While the former often theorises and assumes the capabilities of existing tools, the latter provides practical validation of current applications and capabilities through real-world professional experience. This combined approach, integrating different perspectives, provides a comprehensive view that can serve as a definitive starting point for future exploration. This foundation could help to develop new functionalities and improve products to better meet specific user needs.

Methodology

A multimethod approach was used, with research conducted using a variety of tools: → reviewing academic knowledge from the literature, → conducting targeted research, with direct involvement of professionals through surveys, → exploring perceptions of technology and knowledge shared on the web (collected through digital ethnography). The findings were then cross-referenced to identify overlaps or possible contradictions between information from different sources.

Figure 01_Methods and objectives



- Method
- Main objectives
- Detailed objectives

Literature: selection criteria

A literature review was conducted in order to systematically map the research done in this area, as well as to identify any existing gaps in knowledge. The search was conducted on the Scopus digital library using the following Boolean search query: “**design AND ai AND “image generation”**” (on the 2nd February 2024). Articles were restricted using criteria common in systematic reviews:

- **Year filter: 2020-2024. Articles published within the last 5 years;**
- **Document type: conference papers, articles, reviews, letters were all included (the only excluded typology was “conference reviews”);**
- **Language: only results available in English were included;**

Using the above parameters a total of 50 were gathered. Three of these did not have a full text available for access by the author’s institution (Politecnico di Milano). The articles were reduced to 25 after manual title and abstract filtering and the selected were read in full to determine relevance to the domain. After **snowball sampling** (if a reviewed article cited a relevant sounding article, it was subsequently assessed and, if deemed relevant, added to the pool of articles for review) six articles were added. The final 19 papers constitute the primary analysis of this research.

As described above, only recent articles were considered relevant. However, upon reading, it became clear that the rapid advances in generative models make literature even a few years old less relevant, as described by one of the sources (Mathiesen and Canossa, 2023). The articles selected from the direct search were all published between 2023 and 2024. The only articles from years before 2020 were later included using the snowball method because they were considered particularly relevant to the specific topic addressed, especially when comparing the different applications of artificial intelligence in the design process. To be included in the review, papers needed to be relevant to the topic. The relevancy measure excluded:

- **articles not directly related to the generation of images from text prompts that differ in the type of prompt provided by the user or the result (e.g. image-to-image only, image-to-video, modification of a provided image, or chatbots);**
- **articles focused on the purely technical aspect of machine learning (e.g. model training and image classification) and not on the interaction of the model with the user**
- **articles related to the application of image generation in fields too distant from the one considered in this research (e.g. tourism marketing, car design, robotics, medical field for predicting genes and prognosis from images, indoor air distribution...)**
- **articles interested in the pure comparison of the generated AI product, as a “finished product”, with “human craftsmanship”.**

The literature reviewed includes information from **grey literature** sources. Most grey literature is considered less prestigious, reliable, and “official” than publication in a peer-reviewed journal. It is often used to publish early results or contain information that has not been fully investigated and validated in scientific studies (Cornell University Library, 2021). With mainstream access to text-to-image generative

models, users organised themselves on social media to share their creations and tips. Midjourney users post their requests on public Discord channels, accessible to all. Some specifiers have even gone viral. Sanchez (2023) has acknowledged the prompting as a social process. It is therefore worth highlighting the **diffusion of knowledge through non-scientific channels**, in particular blogs (e.g. Synthesis AI), social platforms and media (e.g. Medium, Hugging Face...), which drive experimentation in text-to-image generation and create communities of enthusiasts who share their findings. It was considered important to take these channels into account in order to synthesise and evaluate the available evidence.

User involvement: method and objectives

The survey was conducted in English (34 responses) and Italian (18 responses), with a total of 52 responses collected. As the survey had two versions, some of the answers to open-ended questions are reported in the original language and then translated. It was distributed through various social channels (one Discord server, one Facebook group, one LinkedIn group, one Reddit community) and through private groups on Signal and Whatsapp applications (the groups included students from the Digital and Interaction Design course at the Politecnico di Milano). See the *Appendix 01* for the list of specific channels.

This research did not target gender or age, so no personal information was collected. The user pool was selected on the basis of professional expertise: all users are designers or design students who have worked with digital interfaces. Questionnaires are a quantitative data collection tool. Given the small sample of professionals involved, **the data collected cannot be considered quantitatively representative**. This factor hinders the generalizability of the research. Future studies could expand the size and diversity of the subjects to address this limitation. The typology of the data itself, which included some open questions, and the sincere and interested participation of almost all the users involved (only 2 users did not answer this type of question), allow them to be analysed and used in a **qualitative approach**. However, the method of distribution allowed the collection of more opinions than a qualitative technique, such as interviews. The survey was conducted to gather data relevant to the specific research that had not been verified by literature review or that needed further investigation and verification. In particular, the questionnaire was constructed on the basis of several sections, each with a specific purpose. The full structure of the questionnaire and the questions to be addressed can be found in the *Appendix 02*.

The first part of the survey aims to gather information about the current design process followed by practitioners. More specific information will then be collected in the context of the particular design process preferred by the user (from the options: "User-Centered Design (UCD)/Human-Centered Design (HCD), Design Thinking, Agile Design, Lean UX, Goal-Directed Design, other..."). The following questions **explore the role of image research in the process**: at what stage is it most useful, what is its relevance, have users ever encountered problems with image research, how have they dealt with them and, most importantly, what are the criteria used to select inspirational images. The study of current selection criteria is relevant to the selection of criteria for qualitative evaluation of generated images

in future tests, and thus generation models. The identification will be compared with the conclusions drawn from the literature review. The aim is to create a standardised evaluation proposal specifically for the purpose of user interface design, based on the needs/requirements and expectations of a practitioner in the field. Furthermore, can the current difficulties in the “classical” research process be overcome by generating original images based on descriptions? A final question directs the user to the next section. It is important to understand the **current popularity and integration of generative AI tools in the UI design process.**

In the second section, questions are raised to **compare the possible role of AI-generated images with conventional image search**, in an attempt to answer the following questions:

- **Are their roles complementary?**
- **Are there problems that can be better addressed and solved by generated AI images?**

Depending on whether integration has already taken place or not, the user is asked about a possible and proposed or current way of integration. If the user has already used them, further questions aim to understand the depth of the exploration: whether there is a preference for a particular tool and why. Some predefined options have been presented in order to facilitate the response and to help the user to think about different aspects related to the results (“Aesthetic result of the generated images”, “Capability to create a diverse range of images, accommodating various design needs”, “Consistency and alignment of the result with the prompt provided”, “Ability to generate high-resolution images”), but also about the ease of the interaction with the tool (“Easy prompt creation using natural language and conversational descriptions”, “Easy prompt construction by comma-separated list of values/keywords (nouns and adjectives)”, “Intuitiveness and ease of use in the tool's interface”), cost and time-efficiency (“Free use”, “Affordability for users with budget constraints”, “Quick processing and generation times for images”) complementary aspects offered by the service, in terms of experience (“Presence of a supportive user community”, “Regular updates and improvements from the tool provider”, “Availability of tutorials, documentation, or educational resources”). This path involves an additional section devoted to understanding how the user creates the prompt (“I independently invent the prompt”, “I do research and follow tips or guides”, “I use other AI tools (like ChatGPT) to generate the prompt”, “other...”), any methodologies developed and whether these vary from one tool to another.

In many cases, the user is given the option to select more than one answer, together with the possibility to add any other answers that were not considered by the questionnaire author. Open-ended questions have been left non-compulsory, leaving the user free to elaborate.

Digital ethnography: methodological boundaries

Online ethnographic methods were employed to gain qualitative insights into virtual communities, by capturing user opinions on the specific topic. The first step in this method was to set a clear research objective with a narrow and applicable question. Selecting data sources for analysis was crucial to meeting this goal. Although Reddit and Discord host the largest amount of content produced on generative AI applications, with dedicated subreddits and servers, the primary focus on these platforms is on sharing results and prompts. These provide limited clues to user perceptions, due to the lack of opinionated comments. An exploratory review of Instagram revealed the platform was unsuitable for detailed analysis because of the minimal content and engagement of users. **YouTube** proved to be the most appropriate channel for this study, offering tutorials that promote practice and information sharing. **The longer format of the content engages users seeking specific knowledge and encourages them to contribute and participate in the comments section, where opinions are exchanged.**

Given the relative scarcity of content also on YouTube, the research employed a qualitative approach, manually scraping data to collect direct quotations. Both user comments and creator contents (from video transcripts) were considered. The selected videos were produced by professionals and regular users who tested the tools and shared their personal views. Therefore, it wasn't considered necessary to distinguish between the reactions of the creators and those of the viewers, which can sometimes be aligned. The examination conducted on textual data aimed at **Thematic Analysis**, to identify and categorise recurring opinions, themes and patterns of ideas related to the topic. For collecting comments, the "YouTube Data Tools - Video Info and Comments Module" was employed, while transcripts were obtained using "youtubetranscript.com".

In the literature review, tools were categorised based on their role and functionalities, leading to the selection of tutorials focused on:

- **Image generators: a total of 12 videos with approximately 3.5k comments, from 5 different channels (Table 01);**
- **UI editor/prototyping tools: a total of 7 videos with 685 comments, from 6 different channels (Table 02).**

For the former, tutorials examining Midjourney and DALL·E 3 were identified. Finding a diverse range of content on image generators proved challenging. The videos have been selected from the results of the search query: **"Midjourney and dalle for UI and web design."** Three videos were added from links referenced in the initial findings. These were also found to be relevant to the consideration of DALL·E 3, which was not initially included despite the search key. The UI editors analysed in the tutorials: Galileo AI, Auto Designer by Uizard, Genius and Framer AI. Searches were conducted using the format **"UI editor name + review"**, for example, "Galileo AI review," "Uizard AutoDesigner review," and "Framer AI review." Exclusions from both searches were:

- **videos sponsored or produced by the tool's creators;**
- **videos with less than 1k views;**
- **content irrelevant to the focus or not matching the search query.**

Much of the excluded content involved the broad application of AI (such as ChatGPT for UX exploration), image generation not intended for UI design, or focused extensively on prompt engineering. Some tutorials covered the entire translation process from image to final layout, with significant sections devoted to non-AI topics. Given the lack of content dedicated to UI and web design, different searches often returned the same results.

Table 01 (right)_Reference list for image generator videos

Table 02 (next page)_Reference list for UI editor videos

| <i>Channel name</i> | <i>Published</i> | <i>Title</i> | <i>Link</i> | <i>Accessed</i> |
|---------------------|------------------|---|---|-----------------|
| Caler Edwards | 17 Jan 2023 | Using AI for UI Design is INSANE! | https://www.youtube.com/watch?v=tov-jr57FjTw&t=12s | 3 Mar 2024 |
| Codex Community | 10 Jan 2023 | Reacting to Beautiful AI Art Web Designs + How to build them | https://www.youtube.com/watch?v=qjllrnvc-EA&t=1013s | 3 Mar 2024 |
| Codex Community | 24 Jan 2023 | How to use AI Art to Redesign Websites with MidJourney | https://www.youtube.com/watch?v=st-FORMbOQS8&t=239s | 3 Mar 2024 |
| Codex Community | 24 Mar 2023 | New AI Art Web Design with MidJourney 5 (v4 versus v5) | https://www.youtube.com/watch?v=gVA3FCqy7lc | 3 Mar 2024 |
| Codex Community | 20 May 2023 | How MidJourney 5.1 brings new changes to Web Designs AI Art | https://www.youtube.com/watch?v=0igwt0djJEI | 3 Mar 2024 |
| Codex Community | 7 Jul 2023 | 10 Web Design Tricks using Midjourney | https://www.youtube.com/watch?v=mE-KEyKw0QLM | 3 Mar 2024 |
| Codex Community | 21 Dec 2022 | AI Art Web Design is Amazing - Are AI Website UI The Future? | https://www.youtube.com/watch?v=fFMSm-mW1LIY | 3 Mar 2024 |
| Codex Community | 27 Dec 2022 | How to use AI Art and ChatGPT to Create a Insane Web Designs | https://www.youtube.com/watch?v=8I3N-TE4cn5s&t=374s | 3 Mar 2024 |
| Darrel Wilson | 20 Dec 2022 | I Generated 80 INSANE AI Web Designs in 1 Hour. (SHOCKING RESULTS!) | https://www.youtube.com/watch?v=kuYw_VYrDSs | 3 Mar 2024 |
| DesignCourse | 16 Mar 2023 | Midjourney Version 5 for Web Design - The Results | https://www.youtube.com/watch?v=Z_f9Rn-69PdM | 3 Mar 2024 |
| DesignCourse | 16 Jun 2023 | I asked Framer AI to Build Me 5 Websites | https://www.youtube.com/watch?v=hn1E3X-4LATM | 3 Mar 2024 |
| Josh Mountain | 31 Oct 2023 | Using DALL·E 3 To Design My Software UI: Does It Work? | https://www.youtube.com/watch?v=KD1V0_Br68E | 3 Mar 2024 |

| <i>Channel name</i> | <i>Published</i> | <i>Title</i> | <i>Link</i> | <i>Accessed</i> |
|---------------------|------------------|---|---|-----------------|
| aaoa Streams | 11 Feb 2023 | Exploring Galileo AI: An AI-Powered Design Tool | https://www.youtube.com/watch?v=v-SA3-Xw2xGs | 3 Mar 2024 |
| DesignerUp | 3 Apr 2023 | Watch AI Turn a Single Text Prompt into Stunning UI Designs in SECONDS! 🤖 Uizard Autodesigner | https://www.youtube.com/watch?v=4ilOvROrMtl | 3 Mar 2024 |
| DesignerUp | 14 Jun 2023 | Uizard's Autodesigner is Here! But Has it Gotten Better or Worse?! | https://www.youtube.com/watch?v=VnezicJbfPM | 3 Mar 2024 |
| Flux Academy | 15 Jun 2023 | Framer Ai is INSANE! | https://www.youtube.com/watch?v=QKUnUrs-DzkQ | 3 Mar 2024 |
| Jesse Showalter | 19 Jun 2023 | Is this the end of UI/UX Designers? Auto Designer by uizard | https://www.youtube.com/watch?v=P-jcP0eXyQQQ | 3 Mar 2024 |
| Kim Alban | 12 Feb 2023 | 2 A.I. Design tools to design faster | https://www.youtube.com/watch?v=kVKg-5DG5Rpl&t=146s | 3 Mar 2024 |
| Nathan Gross | 10 Feb 2023 | Will Galileo AI Replace Me as a Designer? | https://www.youtube.com/watch?v=zzYDevictBA | 3 Mar 2024 |

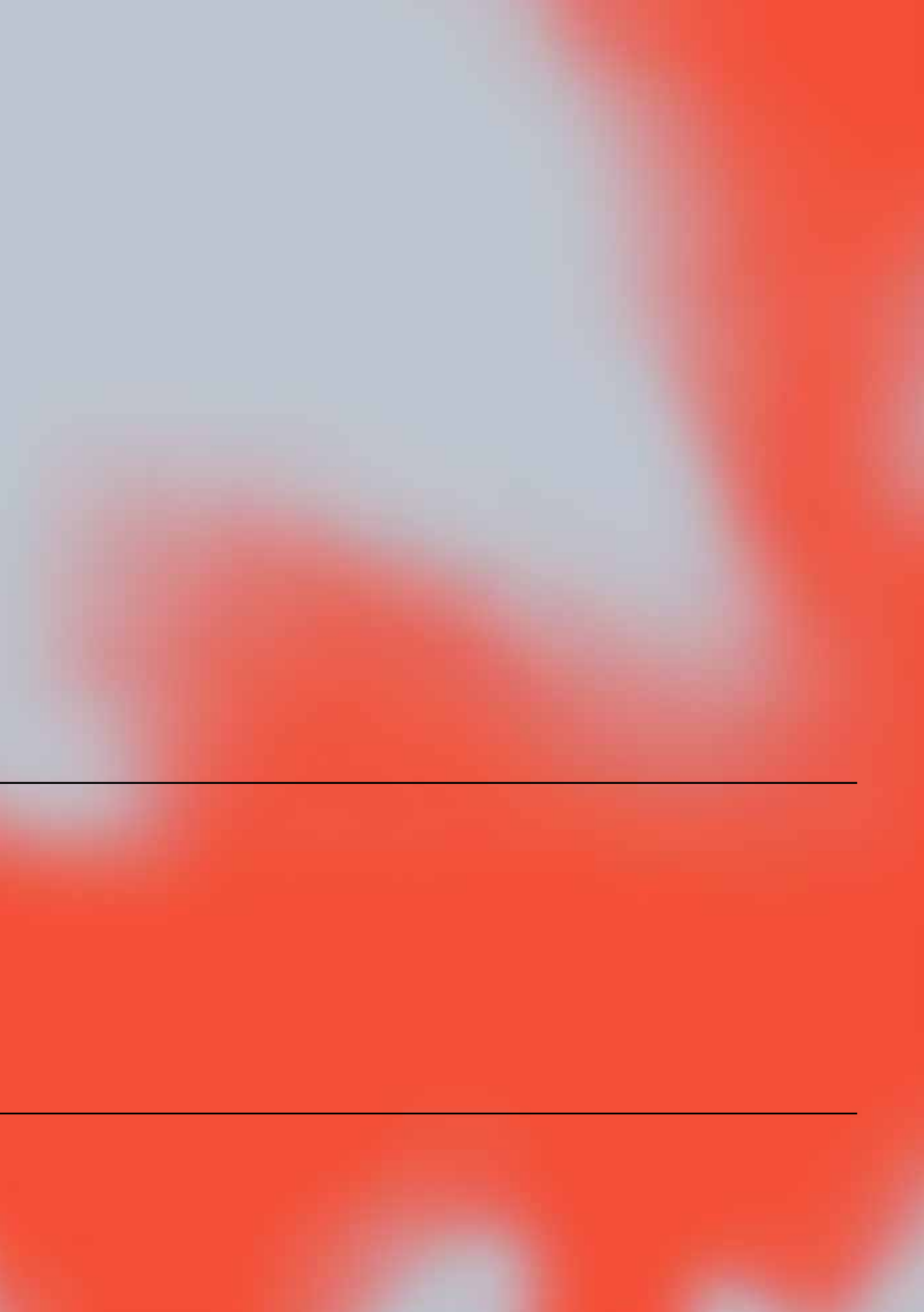
STATE OF THE ART

2.1 Discussion

Literature conclusions
Survey results
Ethnographic research findings

2.2 Triangulation results

Define the role in the design process
Facilitate interaction with the tool
Propose new evaluation criteria



Discussion

Literature conclusions

Research questions

The following research questions were formulated in order to map the current knowledge, the studied applications and the conclusions drawn by the authors:

- What value could image and UI generation bring to the design process? What are the applications of AI that have been specifically experimented within UI design and related fields?
- What challenges do users face in text-to-image generation, and what strategies can be employed to address these challenges?
- How was the qualitative assessment of the results carried out and what criteria were used?
- What are the potential implications of using large models trained on public data when applying the generated images?

1. The role in the design process

The design process is a complex task, requiring imaginative exploration of a broad landscape of possible solutions. A fundamental part of this process involves looking at images and collecting samples to build knowledge about a specific design idea. Research and examination of examples is necessary to assess and learn from what already exists (state of the art and standards) and to borrow elements for your own purposes. In this way, the synthetic design references are embedded within a concept modelling workflow (Stigsen et al., 2023).

Since the advent of the Internet, the most common practice is to conduct image research on one of the various “design gallery platforms”. They support a bottom-up, inspirational process by offering a range of visual examples through the use of a query. Mozaffari et al. (2022) observed that this activity can sometimes result in an overwhelming experience, with possible outcomes including “**design drift**” (the practice of moving design ideas away from the original focus) and **design fixation**. In particular, exposure to a greater number of examples that share common critical characteristics would increase the problem of fixation (Mozaffari et al., 2022). Because of the natural tendency to base ideas on what's familiar,

humans (and therefore designers) focus on a limited set of solutions to a problem. This behaviour can lead to premature convergence on suboptimal solutions rather than broadly exploring a larger space of possibilities (Cai et al., 2023).

New platforms and digital tools are constantly being developed to support the design process. By simulating and modelling design options, these tools can accelerate and enhance cognitive processes and skills. Specifically, **the availability of synthetic image generation services has raised an intriguing alternative to the traditional way of assembling “design vocabularies”**. The existing literature does not provide explicit guidance on how designers should effectively use AI-generated content and the precise roles they should play within the design process (Hanafy, 2023; Wu, Hsiao and Lu, 2024).

However, some experiments have been carried out and various studies have evaluated the influence of AI capabilities on the process. While artificial neural networks are often favoured simply for their ability to **generate fast results**, they also provide us with new feature spaces by assembling unknown patterns and connections based on the data they are trained on (Stigsen et al., 2023). Particularly in application fields such as architectural, product and colour palette design, some studies suggest that professionals can apply AI generative tools in conceptual design to enhance their creativity. In the research conducted by Mozaffari et al. (2022), one of the professionals stated that a beneficial use of AI could be **aggregating vast online information**, including the latest trends in colour palettes and layouts, which individuals might not have the time to search for themselves. This application could serve as a comprehensive reference, merging relevant interfaces to provide insights into current design trends. Hanafy (2023) argued that creating visual representations of design concepts from a textual description (text-to-image generation) could help the architect quickly explore a wide range of possibilities and iterate on different design approaches. Verheijden and Funk’s study (2023) investigated the use of image generation in collaborative and co-design applications through a platform. Although this can be a direct tool for communicating an idea to others, the authors argued that an individual approach to image generation would allow for ad hoc exploration of multiple perspectives without continuous deliberation. Cai et al. (2023) evaluate a baseline condition of image search in comparison to the practice of image generation for inspiration gathering in the context of product design. They conclude that the automatic generation of images from words is <<significantly more inspirational, enjoyable and useful>> than conventional research using popular platforms (in the case of the study, the research was conducted using Pinterest). It may prove insightful for future research to investigate the potential benefits of integrating AI earlier in the design process.

Çiçek, Turhan and Özkar (2023) proposed a different perspective, exploring the potential of the integration of diffusion models into the design education program (in particular during the first year design studio) in order to support students in their first encounters with the ambiguous nature of design problems. Interesting evaluation criteria for the results were selected, taking into account the performance of the AI-generated solution in terms of material properties and abstract concepts representation. The model (in this case only the Midjourney diffusion model was tested) stood out for its **immediate ability to consider and concrete abstract aspects in images**, which were discussed and explained to the students repeatedly during the semester. Lee and Chiu (2023) added observations about the models’ ability to combine stylistic concepts from different worlds. One

of the study participants observed that **the visual stimulus generated by the AI becomes more specific the more detailed the input**. On the contrary, the visual stimulus searched online that matches the keywords of a detailed input is more filtered, making it more difficult to find images. That's why online searches for multiple keyword combinations often yield less than expected results. In contrast, the AI text-to-image generator can easily mix different styles, objects or ways of describing them, "customising" the form and producing unexpected effects. The authors concluded that AI image generation can add complementary and different value to the design process compared to image research. **The two practices shouldn't be mutually exclusive.**

In their research, Lee and Chiu (2023) also observed the different ways in which users approached the tool. They found out that one of the most common practices of users was to have a general direction in mind and describe the texts according to that direction, **expecting the images generated by the AI to match the user's imagination**. However, Stigsen et al. (2023) mention that we actually don't know how all the individual neurons work together to reach the final output (hence, the black-box nature of neural networks). Working with current AI models entails a certain amount of trial and error, in which the user is left to guess and explore ways he could achieve the desired outcome. Typically, they don't produce the required outcomes accurately on the first try. Instead, they ask the user to alter their query language until they get a satisfactory result. This behaviour was also observed by Karahan, Aktaş and Bingöl (2023) where the task was to recreate the picture that most resembled the given one. Participants thought they were more likely to get similar results if they used more words in their descriptions. But in the final choice, they did not always choose what they ended up with, or the picture they produced with more words. The results do not always get better depending on the number of words or how elaborate the prompts are. The way users describe the image can lead the AI-powered tool to many results, and not always the ones they had imagined, so they started to assess their words more carefully.

<<The results of this study show that participants' first impressions of the image and how they express their impressions through description do not correspond with how Midjourney interprets those descriptions.>>
(Karahan, Aktaş and Bingöl, 2023)

This interaction influences communication and opens up different perspectives, which can lead to creative thinking through intended and unintended consequences. Cai et al. (2023) supported the concept of AI image generation as a facilitator of **divergent thinking**. When the designer is searching for visual input, he is not looking for the finished solution. The fashion case study by Cameron and Yanardag is a notable example of the interesting mistakes made by the AI due to the limited training time. The tendency to blend elements from the background into the final design and to combine different types of sleeves within the same dress were found to be inspiring to the pattern makers and led to new ideas (Hughes, Zhu and Bednarz, 2021). Yang et al. (2021) created an icon dataset through application of a generative adversarial network (GAN). Generated icons of **lower quality were found to be more inspiring** by the participating designers, which could actually improve their design effectiveness. Similarly, in Stigsen et al.'s study (2023), the aim was to use AI image generation to develop a focused but still abstract visual vocabulary, as opposed to the "rather incoherent but realistic" Pinterest and Instagram boards. Future research may therefore explore the generation of more abstract visual cues (Yang et al., 2021).

It is interesting to note the results of studies on the development of tools to assist the creation of realistic layouts, such as the one conducted by Mozaffari et al. (2022) where professional designers were involved to investigate the possible value in real design practice. **The ability to provide enough abstraction was highlighted as the most interesting aspect of the generated images.** They were able to convey the design concept while leaving out unnecessary details, thus stimulating creativity. Abstract design examples help increase creative freedom, reducing design fixation. The randomness of the calculation method and the ambiguity of the textual description confer to the AI generated images a unique, fuzzy and customised nature that can meet the needs of designers' reference (Lee and Chiu, 2023). This makes the second approach observed by Lee and Chiu (2023) particularly interesting, according to which **some users do not provide detailed and specific input, wondering what images the AI would generate.**

The literature review identified case studies specifically related to UI design. Although this research is intended to focus on text-to-image technology, it was considered important to mention cases that do not fall entirely within this category. Given the specific application, it was preferred to look at the wider literature to understand which modalities have been explored and why they have been chosen over others. Nguyen et al. (2018) already in 2018 proposed the creation of a tool that, trained with screens of applications and descriptions provided by humans in natural language, was able to identify patterns in the samples of the dataset and recreate the required interface. Despite the fact that only mobile interfaces have been addressed and the pattern observation and labelling process have been done manually, the task of creating the user interface is accomplished by the developer with the sole support of this tool and a textual description. The general idea is to **replace the manual and time-consuming creative process of the UI/UX designer** (Nguyen et al., 2018). However, the study examines the construction of the "login screen". If the tool can resolve graphical elements with a more complex level of interaction, without the intervention of a designer, is still an open issue. Another case study from 2019 by Li et al. proposed to solve the refinement of UI layouts through the accurate alignment of a set of randomly placed 2D elements. The tool automated the technical practice of arranging graphical components to create an aesthetically pleasing and standardised interface, based on similar UI screenshots.

<<The generator has the desirable property of being permutation-invariant: it will generate the same layout if we reorder the input elements.>>
(Li et al., 2019)

It is therefore not a generative tool. And its use cannot be placed in the context of creative inspiration at an early stage of exploration, but it does fulfil the task of producing a realistic layout page. The tool proposed instead by Mozaffari et al. (2022) involved the application of a style-based generative adversarial network (called "StyleGAN"), trained with a dataset of existing interface designs, to generate a "diverse and yet focused" set of examples based on a preliminary design input (e.g. a UI mockup image). One of the problems lied in creating the dataset on which the model was trained. As the data consisted of UI screenshots from the same platform, there were obvious limitations in terms of homogeneity and outdated designs. The authors state that the aim of the tool should be to provide ideas for layout or structural changes, suggestions for alternative component designs and for different aesthetics, such as colour schemes. Given the results of other studies (preference for abstraction and fuzziness of results), it's still unclear whether the

tool can really provide support. The use of large models with access to the huge amount of data available on the web could partially solve this issue and help to produce more differentiated results (although this does not come without problems, including those related to copyright). But the aspect that stands out in these examples is the particularly high level of user control over detail, whether it's linked to predefined settings (<<including background colour, primary font, number of columns and visual density>> (Li et al., 2019)) or a necessary "preliminary design artefact" to be submitted to the tool (Mozaffari et al., 2022). The interaction is initiated by a starting point that needs to be found through research or self-produced. Rather than supporting the design process, these platforms are an **attempt to automate the final step, converging the idea and finalising the layout**. Therefore, the tools play a different role from that introduced by image generation as investigated and reported by other studies, which would not remove its utility for other purposes. They could be easily applied in a context that requires the rapid delivery of graphic layouts for a large customer base, with a different positioning in the market.

The literature is divided on whether these technologies should be designed to support or automate creative practice and, more specifically, user interface (UI) design. The inner workings of the algorithms often remain a mystery to the user and even to the model's engineers. This opacity can lead to a fundamental mistrust between end users and the systems they interact with. Mistrust often goes hand in hand with fear of the new technology, in particular of being replaced by it. It's necessary to explore and develop a relationship with the underlying system to understand its characteristics, capabilities and limitations (Hughes, Zhu and Bednarz, 2021). For the acceptance and the adoption of a new tool by humans, and so by designers, there are requirements to take into consideration: users need to feel in control and understand the consequences of the assistance; they should be given enough time to learn their own way of using the tool and be allowed to iteratively refine problems and solutions (Hegemann et al., 2023). Lee and Chiu (2023) stated that computational systems are creative.

<<[...] the form of their creativity is not the same as we think and will never be exactly the same as humans>> (Lee and Chiu, 2023)

Humans are not good at repetitive work and handling large amounts of data. But AI can't replicate human experience, so the two different creative forces are not, or should not be, in competition.

<<efforts to augment human creativity with computational tools are much more likely to be successful in the foreseeable future than attempts to completely replace human creativity [...] if we don't focus on developing such synergistic human-computer systems, we - as a society - may fail to benefit from their potential.>> (Cai et al., 2023)

The authors explored the problem of "the desirable task allocation" in human-computer systems and defined the ability to perform a task in collaboration as a "specialised collective intelligence". In a collaborative approach, the most critical aspect is how the division of tasks to be performed takes place. It is argued that the development of artificial intelligence is gradually shifting the design process from "designer creation with machine assistance" to "**machine creation with designer evaluation**." (Wu, Hsiao and Lu, 2024). Assisted creation systems have evolved across three distinct generations, increasingly making creative processes

more accessible. Initially, these systems offered digital means with limited support, but as they were provided with “greater agency”, they achieved the capability to engage in meaningful dialogue. This is characterised by robust action-feedback loops, where control is jointly held, and decisions are collaboratively made with the user (Pieters and Wingers, 2016). In the context of text-to-image technology, when prompting an image generator with the description of a product, it doesn’t “design” it. It simply generates an image that can be a very high-quality flawless representation of that product. The next time, a different shape of the same product will be generated, in random variation. No conscious decisions are made to create coherence between the form, the object’s function, and the interaction with it (Verheijden and Funk, 2023).

Çiçek, Turhan and Özkar (2023) also evaluated the model in terms of its ability to link the various generated solutions throughout the design process, defined by sequential briefs and stages. In this dimension in particular, it received lower scores. This indicates that the evaluated diffusion model is not yet able to perform a continuous workflow that creates a proper “logical connection” between solutions. Nor is it able to mimic intentionality and critical judgement. Therefore, **the implementation of AI tools in a formal setting would not disregard the importance of design reasoning mechanisms**. The entire design process is a multi-faceted work, that has to take into account functionality, user experience, aesthetics, and engineering, to solve the sorted out problems. It doesn’t rely just on creativity but also on **non-linear thinking** (Lee and Chiu, 2023).

Moreover, the designer role is not limited to solving the problem, but also to identifying it (problem finding), evaluating it and defining action plans to deal with it (problem setting). It enables a differentiation of solutions that meet the specific needs of the targeted user. Karahan, Aktaş and Bingöl (2023) highlighted that, despite the ability of AI tools to generate images in seconds, the design process characteristic “non-linearity” remains. The tool’s suggestions may change the designer’s initial direction and provide possible new solution ideas. However, the ongoing act of evaluation remains in the hands of the designer. Even in the case of attempts at automation, it could be argued that the figure of the developer fulfils the role generally assigned to the designer in deciding whether or not to apply the generated layout. This would suggest a shift and a reassessment of the division of labour, but not a real abolition of the role.

2. Interaction with the tools

When investigating how to interact with image generation tools, examples related to the construction and analysis of text prompts were considered. Since text-to-image generation is based on the information contained in the text prompt, the words chosen play a fundamental role in determining the result. Different descriptions can emphasise different aspects of the image, leading to unexpected results. The first aspect to consider when interacting with the tool is the **inherent limitations** of the model that affect interaction and prompt construction. Among the most common are those mentioned by Sanchez (2023):

- **artifacting: creation of unrealistic features, mainly affecting hands and faces, or due to the presence of watermarks from stock images in the training set;**
- **concept entanglement: unintended associations within some**

keywords of the prompt.

As Sanchez (2023) found in their study, the difficulty of formulating prompts is one of the main challenges faced by users. But this is often seen as a communication issue or an obstacle independent from the system. A growing practice, and one way to overcome the problem, is **to use other AI tools** (particularly chatbots) **to build the textual input for image generation**. The exploration of Wu, Hsiao and Lu (2024) represents an interesting case identified in the literature for the use of GPT-4.0 as Midjourney's prompt generator. The GPT-4.0 was required to act with specific roles (e.g. senior product designer or Midjourney prompt creator), providing him with the necessary design background information and some prompt examples of product forms (to let him learn the expression paradigm). Finally, GPT-4.0 was asked to generate prompts.

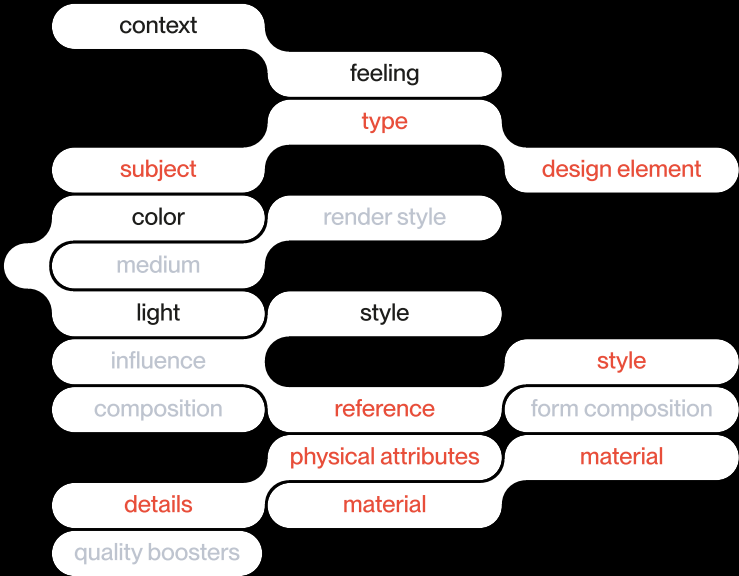
Some studies support prompt engineering by identifying categories of words to include in the text prompt and suggesting a structure to achieve more satisfactory results. By observing a set of real user text prompts directed at the Stable Diffusion V1 model (from the open source dataset DiffusionDB), Sanchez (2023) identified semantic structural patterns and common practices (such as repetitions and specific word combinations, also called "prompt specifiers"). The results collected in the study were used to construct a **prompt specifier taxonomy** in order to design more meaningful interactions with generators. The developed specifier taxonomy identifies the following semantic categories:

- **Context:** attributes (usually added before the subject) that provide contextual information and cannot be fully represented, such as the era, the weather, and emotions;
- **Color:** global colour schemes (also usually added ahead of the subject);
- **Subject:** it refers to elements that are the main focus of the prompt;
- **Medium:** artistic media (photography, painting, rendering, illustration);
- **Influence:** artistic reference (artist names, art genres and movements, artwork names, and art repositories); these tend to be mentioned in sequence, with one artistic reference following another;
- **Light:** lighting conditions in the scene;
- **Composition:** overall arrangement of elements in the image;
- **Detail or quality boosters** (also often mentioned in groups or sequences).

Hanafy (2023) also conducted a query analysis (examining approximately 40 million publicly visible queries from the Midjourney Discord server) specifically related to the architectural domain. They pointed to the common practice of referencing well-known architects, given the instantly recognisable style they can bring to the results. This type of keyword could be likened to the "Influence" category mentioned earlier, supporting the idea that **citing proper terms** in certain areas can help direct AI models more easily and **produce outputs that are conceptually and stylistically accurate** (Hanafy, 2023). Other studies involved direct user testing. In order to investigate how the descriptions provided by the participants would differ, and to code all the information provided by writing text, Karahan, Aktaş and Bingöl (2023) devised six categories: type, style, material, feeling, reference and render style. It's interesting to note that the category "**physical attributes**" had to be added after the experiments. In fact, the participants mostly used words related to tangible features in their textual descriptions. However, the

Figure 02_Prompt specifier taxonomies

*Sanchez (2023) *Karahan, Aktaş, Bingöl (2023) *Lee and Chiu (2023)



● Specific vocabulary

○ Secondary vocabulary

authors argue that **this typology of specifiers is not sufficient to achieve desired results**, as abstract concepts such as feelings, style and references help to elaborate textual definitions.

In Lee and Chiu's research (2023), participants had to choose words from given categories of shape description to create their prompts, with the possibility of integrating them with their own suggestions. More than half (four out of six) of the users tended to fill in the elements of style, form composition, material and design elements equally, and only two of them made specific attempts in the text to integrate different style words together. It can be inferred that providing users with specific categories to refer to may influence and even limit their ability to create a more diverse combination of words, when creating the prompt. This behaviour could prevent the generation of more interesting and unexpected results, given the ability of generators to combine and concretise abstract concepts. Stigsen et al. (2023) suggested some alternative and differing ideas for experimentation with prompts. Firstly, they proposed to combine two different vocabularies:

- a specific vocabulary based on provided architectural readings, a foundation of knowledge of basic design principles and historical context;
- a secondary vocabulary, stylistically related, compiled on a topic of the students' choice.

The choice of words for each prompt had to maintain a balance between architecturally identifiable elements and abstract elements with architectural qualities. In addition, a strategy was developed **using constant words to be kept throughout and variables as words to be interchanged**. This method allowed **maintaining stylistic consistency between the different visualisations generated** (Stigsen et al., 2023). Another interesting finding from the Cai et al. (2023) study is the use of semantically similar nouns to achieve conceptually similar images with very unique stylistic interpretations. Rather than using the same word over and over again, greater variety and success can be achieved by using synonyms (<<"couch", "loveseat", "sofa" [...]>>). Similarly using varying sentence structure led to more varied output.

The mentioned studies are all different in terms of topic, subject and desired outcome (e.g. architectural representations, product design, product colour matching design). In some cases, this led to the identification of specific taxonomies to be included, such as the category "material", which is fundamental when applied to the physical world and becomes unusual in digital design. But it is possible to identify transversal categories that are common and necessary in all fields. Providing the user with guidance and tips to help them communicate more effectively with the tool can improve the experience and interaction, and ensure a higher level of satisfaction. However, given the natural tendency to follow a presented structure, it is important to avoid making these supports limiting, precisely because what text-to-image generation can offer is extreme flexibility. As argued above, **assistive technologies should not homogenise results, but promote discoverability by adapting to users' intentions**, not the other way round (Sanchez, 2023).

3. Evaluation criteria

A key obstacle in this research field is in the assessment of the outcomes obtained.

Evaluating the success of text-to-image models' results is notoriously difficult. In particular, purely **quantitative metrics** (e.g. Fréchet Inception Distance or FID, for quantifying the realism of image generated, and CLIP score, to evaluate how effectively the model makes accurate embeddings) are not sufficient to state whether a text-to-image model can generate "good" images. Especially because it is difficult to translate what "good" means into a numerical value. Moreover, correlations between human and automated metrics are proven to be generally weak. While these metrics **provide valuable insights, they may not effectively capture the nuances of human perception** (Lee et al., 2023). There are several benchmarks for evaluating the performance of state-of-the-art computer vision models. Pascal VOC data sets, ImageNet, SUN database and MS Coco (Microsoft Common Objects in Context dataset) are examples of image datasets with the main aim of object class recognition. Instead, DrawBench (introduced by Imagen) and PartiPrompts (introduced by Parti) are prompt datasets that can be used for qualitative benchmarking.

Several studies point out that little research actually takes human perception and judgement into account, lacking qualitative data. User studies can contribute to a more user-centric and robust assessment. Qualitative evaluation consists in the comparison of text-to-image models to each other according to human preferences. However, **the reliance on human raters introduces subjectivity and the lack of standardisation makes direct model comparison challenging**. Many performed user studies have a similar setup, but there are currently no standard guidelines for their structure. Various papers adopt different fundamental factors such as number of samples, users, models and evaluation metrics (Frolov et al., 2021).

Existing evaluations primarily focus on the consideration of two metrics: **image-text alignment** (also called caption alignment) and **image fidelity** (or sample quality) (Saharia et al., 2022). They tend to overlook other critical aspects. One of the challenges facing current evaluation methods is how to **capture the increasing complexity of results and advances in models**. These models can generate visually striking images and have found **applications in wide-ranging domains** (such as art and design), making current metrics less reliable. Consequently, there is an urgent need for continuous adaptation, expansion of benchmarks and a broader and more diverse set of evaluation criteria. The HEIM benchmark (Holistic Evaluation of Text-To-Image Models) provides a starting point and a new perspective on evaluation. It employs automated metrics in conjunction with human-based ones and introduces new criteria to evaluate the overall aesthetics and originality of landing pages (Goel, 2023; Lee et al., 2023). What follows is a summary of the criteria that were taken into account in the reviewed literature and how they were selected, with the aim of identifying those most relevant to the intended application. Due to the integration of AI generation technology in the field of design, **the evaluation metrics must respond to the needs of the professionals' specific target audience**.

Some recently developed design inspiration tools use algorithms to determine **image similarity**. The result of the interaction is a list of suggestions similar to the user's input. The choice of this approach is related to the need for the proposed examples to be relevant to the subject of the research. However, relevance and usefulness differ. Greater similarity may be appropriate in a top-down process, useful when the designer has a concrete idea in mind and is looking for examples to implement it. Yet **it can lead to potential design fixation and hinder the creative**

process, especially in the early stages (Mozaffari et al., 2022). Mozaffari et al. (2022) suggested a distinction between the necessary similarity of the outputs to the input image (in this study, the user input is image-based) and the diversity within the set of outputs. Participants were asked to provide direct feedback on design examples to assess the ability of a set of UI images to support inspiration, based on the following criteria:

- their relevance to the design task
- the diversity of the design examples
- the effectiveness of the examples for inspiration
- the general perception (positive and negative)

Cai et al. (2023) proposed similar metrics to evaluate the image generation compared to image search activity, including concepts like usefulness ("Using this tool would enhance my effectiveness in my job") and **inspiration** ("I experienced inspiration during this task"), but introducing the concept of the enjoyment of the activity ("I liked doing the activity"). Users were asked to rate their subjective experience on a scale of 1 (strongly disagree) to 7 (strongly agree). The time taken to complete the activity and the number of interactions required to achieve a satisfactory result were also included in the rating.

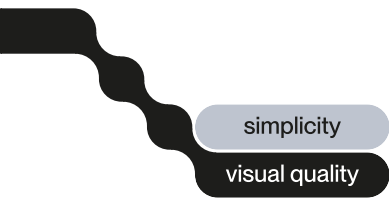
Qualitative comments and feedback could lead to the consideration of adding the concept of **surprise** as a possible metric to define the **ability of the model to generate unexpected results**. The actual benefits of the level of diversity of outcomes (high or low) have been explored, with no clear value in terms of inspiration



for a high diversity approach. The issue of diversity is an argument that could be extensively debated and researched. As discussed by Cai et al. (2023), **it's not clear whether this kind of technology could really support convergence**, even by engineering the prompt to achieve more specific generations. In addition, the results of the study indicated that the **visual quality** of the generated images had an impact on how well designers perceived the examples. This supports the argument that quality should always be used as a metric for evaluation.

The assessment criteria chosen in Çiçek, Turhan and Özkar (2023) study were strictly related to the educational context in which the AI tool was applied, rather than the functional application in a working project. They included the overall success of the final result (in the case of research, to the given assignment), but also distinguished the performance of the solution in terms of material properties and abstract concept representation. Finally, they took into account the ability to connect the solutions generated by the sequential tasks and stages of the process.

Yang et al. (2021) combined Apple's icon design requirements and suggestions for their evaluation process. Five subjective criteria were selected: legibility, simplicity, uniqueness, artistry and attraction. The authors took into account the usefulness principle, also mentioned in other studies, by adding the open-ended qualitative question at the end of the questionnaire: "How helpful do you think the generated icons in this study are to your daily icon design work?". Given the sectoral theme, which relates exclusively to icons, characteristics such as legibility and simplicity are relevant. On the other hand, the principles of **uniqueness** and **artistry** are interesting in a more general sense, as they can be linked to the idea of surprise,



*Çiçek, Turhan and Özkar (2023)

*Lee et al. (2023)

*Mozaffari et al. (2022)

*Yang et al. (2021)

*Cai et al. (2023)

Figure 03_ Comparing criteria used in different studies

unexpectedness of results and diversity, as if to **assess the originality of the generation**. The attraction metric can instead be linked to the general perception of the generated image, likely influenced by the aesthetics of the results.

4. Ethical concerns

Possible plagiarism issues that may arise from taking inspiration and referencing online resources is a factor to consider when developing new proposals. One participant in Lee and Chiu's study (2023) stated:

<<I would feel more comfortable in extracting the design elements of the AI-generated visual stimulation [...] there is no problem of accidentally drawing too much [...]>>

AI's ability to always create new and unique content is given by the production of probabilistic results, not predetermined outcomes. In fact, **it is impossible to create exactly the same image twice**, even with the same prompt (Siddiqui, 2023). However, it's important to be aware of the ethical concerns surrounding the provenance of data used to train deep learning models. As the mega-datasets (including all publicly available material) are too large for proper human curation, **the intellectual origins of the data are largely unknown**. This, combined with the ability of trained models to memorise their training data, rather than learn to generalise from their structure, raises concerns about the originality of outputs and **copyright infringement**. (Stigsen et al., 2023; Kenthapadi, Lakkaraju and Rajani, 2023). If the model fails to produce new original results, it increases the margins of copyright violation of the used training data. It is therefore important to define metrics for **evaluating AI-generated images in terms of diversity and uniqueness to prove the ability of the model to produce images that are novel** (Abdallah and Estévez, 2023).

The use of AI-generated images as "final products" poses challenges, especially when prompts incorporate specific stylistic and aesthetic references through artist names, repositories or quality boosters. Using them as a source of inspiration may be a more valuable and trustworthy application, acknowledging the individuality of perception and interpretation of the same visual stimuli by different individuals.

Summary

The constant development of digital platforms and tools supports the evolving design process, with AI-generated synthetic images emerging as an intriguing alternative. However, current literature **lacks clear directives for designers on the effective incorporation of AI-generated content within their workflow** (Hanafy, 2023; Wu, Hsiao and Lu, 2024).

Interdisciplinary studies indicate that AI tools support creativity in the conceptual phase of architecture (Stigsen et al., 2023; Hanafy, 2023), product (Cai et al., 2023), and colour palette design (Wu, Hsiao and Lu, 2024) by **promoting divergent thinking**. Yet, these tools seldom deliver accurate results on initial attempts. The descriptive prompt by users can yield diverse outcomes, influencing the communication and thereby **opening possibilities of creative exploration through both intended and unintended results** (Karahan, Aktaş and Bingöl, 2023). Interestingly, some practitioners deliberately provide vague inputs to AI, curious about the unforeseen images it might produce. AI offers the potential to **rapidly** generate diverse options and **iterate** various design approaches (Hanafy, 2023). Moreover, its ability to **materialise abstract ideas** (Çiçek, Turhan and Özkar, 2023) and **meld stylistic elements** (Lee and Chiu, 2023) assembles unique, unpredictable patterns and connections (Stigsen et al., 2023). Its inherent randomness and the ambiguity of the textual descriptions result in distinctive images that possess a valued fuzziness. They are able to convey design concepts without unnecessary detail, stimulating creative thought (Lee and Chiu, 2023).

Employing AI-generated images as “final products” raises ethical considerations concerning the intellectual origins of the vast datasets used to train deep learning models, which are too large for proper human curation. The potential for these models to replicate rather than interpret the data structure they're trained on further complicates matters, introducing **concerns about the originality of the generated content and potential copyright issues** (Stigsen et al., 2023; Kenthapadi, Lakkaraju and Rajani, 2023). A more strategic application of AI in design may involve its early integration in the design process.

In contrast to traditional image research, AI-generated images offer a complementary approach: detailed prompts lead to more filtered search results online, whereas **AI combines elements in an additive way**. While the creation of images via AI is constrained by a maximum limit, gallery platforms often offer an extensive array of results that may seem overwhelming due to their similarity (Lee and Chiu, 2023).

Recent advancements propose UI design tools that suggest examples based on different inputs (like wireframes and existing designs), employing algorithms to determine image similarity or refine layouts (Nguyen et al., 2018; Li et al., 2019). Designed to **automate the final design steps**, these tools enable rapid delivery to a wide range of customers without diminishing the role of image creation in other creative explorations.

The design paradigm appears to be shifting towards a "**machine creation with designer evaluation**" model, where computational tools should be envisioned to augment human creativity through dynamic human-computer collaboration, (Wu, Hsiao and Lu, 2024) **without replicating the intentionality and critical judgement** inherent in human designers. Hence, integrating AI tools into formal design settings

should complement rather than replace traditional design reasoning mechanisms.

One of the significant obstacles in using AI for image generation is the challenge of crafting effective prompts (Sanchez, 2023). To mitigate this, an emerging solution involves **leveraging other AI tools**, such as ChatGPT-4.0, which can fulfill specific roles and serve **as a prompt generator** (Wu, Hsiao and Lu, 2024). Research highlights the benefit of "prompt engineering", introducing a **taxonomy of "prompt specifiers"** (Sanchez, 2023) to guide users towards achieving more precise and conceptually coherent outputs, including the strategy of referencing specific styles for instantly recognizable results (Hanafy, 2023). Achieving the desired outcomes goes beyond simply describing physical attributes; **incorporating abstract elements like emotional tones and stylistic nuances** into prompts enhances the richness of the generated images (Karahan, Aktaş and Bingöl, 2023). Additionally, employing semantically similar terms and varying sentence structures can lead to conceptually similar yet stylistically distinct results (Cai et al., 2023). Despite these strategies, it's crucial to preserve the inherent flexibility of text-to-image AI in providing instructions to users. **Over-reliance on structured approaches potentially constrains the creative possibilities** by directing users too narrowly and confining them to predetermined outcomes (Lee and Chiu, 2023). Categories and structured guidance should suggest ideas and not restrict them, fostering a wide range of explorations that reflect the user's intent and beyond (Sanchez, 2023).

There are currently **no standardised guidelines for qualitative assessments** in AI-generated imagery. Existing evaluations primarily focus on two metrics: image-text alignment and image fidelity, which measure the semantic correlation between description and image, and the realism of the generated image, respectively (Saharia et al., 2022). Studies have emphasised that the visual quality and aesthetics of images significantly influence designer perception (Cai et al., 2023). However, with the evolving complexity of AI models and their **integration into design fields**, there's a pressing need to develop a broader and more diverse **set of evaluation criteria tailored to professional needs** (Lee et al., 2023).

The **relevance** of generated images to **project themes** is crucial, yet the potential for design fixation due to overly similar outputs poses a risk to creativity, particularly in the initial stages of design. It's crucial to measure both the similarity with the initial input and the diversity within the set of generated results (Mozaffari et al., 2022). Furthermore, given that neither relevance nor similarity necessarily equate to the **effectiveness of the images in inspiring** the design process, the metric should be introduced to assess the actual support the images can provide to professionals. Feedback from participants has pointed to the unexpected nature of results and the resultant positive surprise as valuable (Cai et al., 2023), indicating the potential for metrics assessing **originality and uniqueness**, as already seen in the HEIM benchmark (Goel, 2023; Lee et al., 2023). The criteria can also evaluate the ability of the AI model to produce images that are novel, avoiding the risk of copyright infringement.

Survey results

Image research: role, importance and criteria

Nearly all designers, except for nine, **use image research as an inspiration source during the ideation phase**. Additionally, many continue to seek visual inspiration into the development phase (34 designers), specifically when crafting solutions, while only 15 utilise this research for creating detailed presentation images. Image research emerges as a critical component in the design process, with nearly 40 designers rating its importance as 3 or 4 on a 4-point scale, underscoring its significant role across various stages of design.

Designers were asked about the typical design methodology employed in their projects to explore the potential link with the conducted image research, aiming to contextualise the role of visual references in their processes. The small sample size makes it challenging to establish a correlation. Nevertheless, the data indicates that almost all designers following User-centered and Design Thinking approaches incorporate image research during the initial ideation and inspiration stages. However, those following a User-centered methodology continue to leverage image research into the development phase, while less than half of those practicing Design Thinking do so. This difference could be attributed to Design Thinking's focus on innovation and developing new solutions through direct user testing from an early MVP stage. User-centered design, on the other hand, seeks to improve usability and user experience, which may lead designers to engage more thoroughly in ongoing research and benchmarking to identify standards and the state of the art.

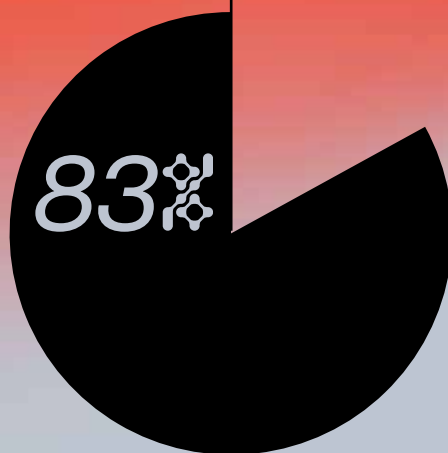
Participants were then asked about their methods for selecting inspirational images for UI design projects. The results highlighted the critical importance of the style of the references, specifically "**Consistency with brand identity and style**," and their alignment with the project's theme or subject matter, with almost 40 respondents prioritising these criteria. However, these metrics are not the only factors considered: "**Originality and uniqueness of UI elements**", "**Visual quality and resolution**" and "**Colour palette**" were also important considerations for almost half of the respondents, with 27, 24 and 21 users choosing these aspects respectively. Hence, images serve a key role in suggesting "original" design solutions for specific components, with visual quality significantly influencing selections, but possibly for moodboard creation. Although colour palettes can be developed using ad hoc tools, they are often derived from inspirational images. The colours within an image possess a unique capacity to capture or divert our attention, underlining the integral role of colour in engaging viewers and influencing design decisions. On the contrary, the impact of current trends, adherence to conventional layout and UX principles, and the versatility of solutions across various devices appear to be less critical factors.

Surprisingly, more than half (35 respondents) reported **difficulties in locating suitable inspirational images** for their projects, highlighting a potential gap in the availability or discoverability of resources that align with specific design requirements or creative visions. The challenges were attributed to various factors:

Figure 04_Insights
into the use and
application of online
image research

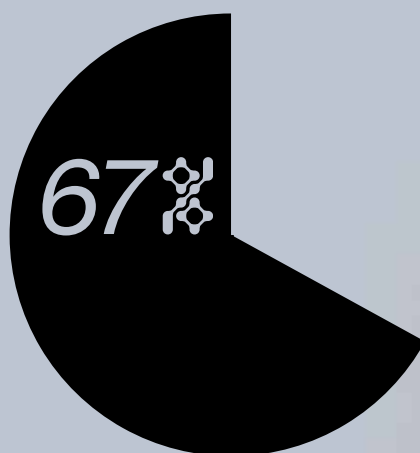
**Designers who use image research
as a source of inspiration during the
ideation phase.**

43/52



**Designers who have encountered
difficulties in finding the right
inspirational images.**

35/52



- **Difficulty in formulating effective search queries (mentioned by 7 users)**
 - <<È difficile prima di tutto capire quali parole chiave utilizzare affinché il motore di ricerca possa comprendere la mia richiesta [...]>> (It is difficult to understand which keywords to use so that the search engine understands my request.)
 - <<[...] se ho in mente un qualcosa di specifico faccio fatica a capire come chiederlo a google/pinterest.>> (If I have something specific in mind, I struggle to understand how to ask Google/Pinterest for it.)
- **Challenges in finding specific results (mentioned by 4 users)**
 - <<It's always hard to find the exact image or concept I have in mind in someone else's images.>>
- **Issues with searching for abstract concepts (mentioned by 2 users)**
 - <<Le immagini non rispecchiavano l'identità del progetto o non riuscivano a concretizzare le idee ed i concept del progetto>> (The images did not reflect the identity or did not materialise the ideas and concepts of the project.)
- **Overwhelming number of results (mentioned by 1 user)**
 - <<There are too many choices which makes it difficult to find out the expected ones. Always looking for better ones.>>

Even in traditional online image search, defining effective keywords is a challenge. Designers often struggle to search for abstract concepts or specific visualisations that truly capture the essence of their design ideas. Furthermore, the vast array of results and choices can overwhelm users, leaving them in a perpetual search for a better alternative. Finally, one participant highlighted their approach to image research, aiming to find a match on an "idea," capture a specific "vibe," or gather inspiration in a more abstract manner. This approach involves **taking inspiration from other domains rather than directly referencing online projects**, and results from an attempt to **avoid any plagiarism issues** that might arise.

Designers' perception of AI

Almost half of the survey participants (25 out of 52) have **not yet incorporated AI-based image generation into their design workflow**, many of whom have never used it. This section aimed to capture the perception of this user group. When asked at what stage AI tools could be useful, 15 out of 25 participants felt that AI could help in the ideation phase. Similarly, a near equal number saw potential in generating detailed presentation imagery, while only nine respondents saw AI as useful in defining possible solutions. In the "Other" option, one participant specified:

<<Considerando che utilizzo la ricerca di immagini come ispirazione, non credo che l'AI possa essere utile in questo use case.>> (Given that I use image search as an inspiration, I don't think AI can be useful in this use case.)

Given the small sample size, drawing definitive conclusions is difficult. However, it's noteworthy that 40% of respondents **doubted the utility of imagery for inspiration, and an even smaller percentage believed it could aid in solution development**. However, another participant stated:

<<Trovo che al momento sia difficile ottenere dei risultati in grado di mostrare bene l'idea di un progetto. È difficile spiegare a parole quello che si ha in mente e quindi ancor più difficile farlo capire all'AI.>> (I find it difficult at the moment to get results that show the idea of a project well. It is difficult to explain in words what you have in mind, and therefore even more difficult to get the AI to understand it.)

This could be in contrast to the idea of possible use at the end of the process for the creation of detailed presentation images. However, this does not mean that the user considered the possible application at another stage of the project (as they did not select more than one option). It emphasises that **some designers may have a specific application idea that is not yet fully successful, and therefore have not yet applied the technology to their projects**. One of the respondents added the use case:

<<To produce results when a work must be done quickly and superficial>>

This would highlight the idea of applying **rapid image production to fast prototyping**. In this user group, "Product images" (17 selections) and "Promotional illustrations" (16 selections) were identified as the graphic elements most effectively created with AI. Only nine participants saw "UI design and digital interfaces" as a viable AI application. This preference underscores a correlation with one of their most suited application ideas, which is the creation of detailed presentation imagery, suggesting that **these users view AI primarily as a tool for enhancing visual content rather than for designing user interfaces**.

Designers' current use of AI

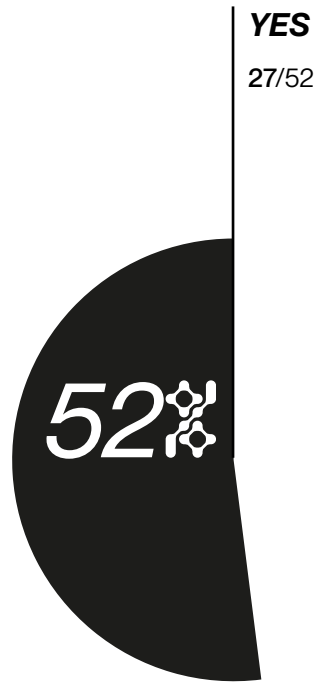
This section explores the integration of AI in the design workflow. Among the 27 designers who utilise AI, 18 employ it for inspiration, 14 for solution development, and just 8 for crafting detailed presentation images during the definition phase. This trend indicates a **preference for using AI earlier in the design process rather than towards its completion**. While only seven designers apply AI directly to UI screens and digital interfaces, this suggests a broader use for generating abstract inspirational images or moodboards.

The most popular AI image generator in this specific user group is Midjourney, with 22 users, followed by DALL·E 3, which is utilised through Bing Image Creator (12) and ChatGPT 4 (11). Additional tools that were selected or mentioned include:

- DALL·E 2 (8)
- Adobe Firefly (7)
- Stable Diffusion model, via DreamStudio API or locally hosted (5)
- Runway (2)
- Leonardo (2)
- DALL·E mini, recently renamed Craiyon (1)

Among the designers who have a preferred AI tool (10 of them), Midjourney stands out as the favourite for 6 users, while DALL·E 3 is preferred by 2 users. The survey asked users to identify the reasons behind their tool preferences, offering insights into the evolving needs and expectations of professionals in the field. The **aesthetic result of the generated images** ranks as the top priority for users (14

Figure 05_ Have you already integrated (or tried to integrate) AI-based image generation tools into your design workflow?



- As a source of inspiration during the ideation phase
- Designing solutions
- In the definition phase, specifically for crafting detailed presentation images



Figure 06_ At which stage of your design process have you utilized AI-based image generation?

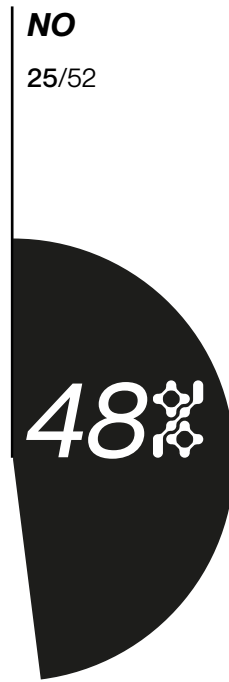
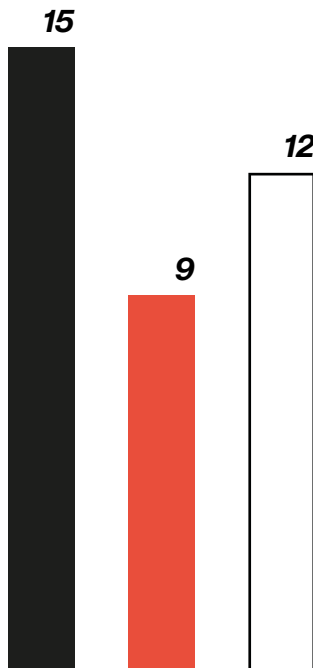


Figure 07_ At what stage do you envision AI tools could be beneficial?



mentions), closely followed by **how well the outcomes match the input prompts** (13 mentions) and the capability to **produce high-resolution images** (13 mentions). This suggests that the desire to use image generation apps is closely tied to these attributes. But, the **versatility in producing a diverse range of images to suit different design requirements** was also considered important and was highlighted 12 times. Cost considerations played a significant role, as evidenced by 13 respondents prioritising "Free use" and 8 highlighting "Affordability for users with limited budgets." This inclination supports the preference for Bing Image Generator, which offers public access, though it contrasts with the selection of paid services like DALL·E 3 via ChatGPT 4 and Midjourney. From an interaction point of view, the simplicity of **natural language and conversational descriptions is preferred** to the comma-separated list of keywords with 12 and 9 responses respectively. **The intuitiveness and ease of use of the tool's interface is not considered a decisive factor** (5). Unfortunately, the users contacted show a lack of enthusiasm for engaging in community activities. Yet, the exchange of practices and tips could potentially foster more satisfying interactions.

Feedback from one participant highlighted that designing a user interface (UI) goes beyond just arranging graphical elements aesthetically; it involves creating a comprehensive experience that addresses the users' needs. Design complexity such as the usability of the interface, the logical flow of interactions, the clarity of the semantics used, and other complex aspects are beyond the full scope and development capabilities of image generators. These aspects require a nuanced understanding and **human-centred approach that AI tools currently cannot fully replicate**, emphasising the need for professional input and oversight to create a coherent and user-friendly design experience.

In addition, another user emphasised that the **output of AI image generation tools "is never the final result"**, pointing out that the iteration and editing process, often facilitated by graphic editors, is essential to fully realise the designer's vision.

Prompt creation and methodology

The questions about prompt modalities in the survey focus on understanding existing practices that can assist designers achieve better interactions and results with AI imaging tools. This approach is aimed at **identifying practical strategies to overcome issues arising from unfamiliarity with these tools**, rather than refining prompt engineering techniques. It's important to note that this research does not imply the author's endorsement of text-based interactions as the optimal user experience (UX) method, nor does it predict this as a future trajectory.

Among the users surveyed, 21 individuals reported creating text prompts on their own, yet only 11 relied solely on this method, indicating a preference for experimentation by mixing various approaches. Fourteen users look up tips or guides, and six consult other AI tools, like ChatGPT, to formulate their prompts. Despite the limited sample size, which may not fully represent the broader user base, it's evident that over half (15 out of 27) lack a clear strategy for engaging with AI image generation tools, suggesting a significant portion of **users navigate these platforms through a process of trial and error**. Several participants shared their strategies for crafting effective prompts for AI tools. Key suggestions included:

→ **Utilising simple, short sentences and clear descriptions or keywords,**

- avoiding complex sentence structures;
- Prioritising features in the prompt based on their desired prominence in the output, potentially repeating important elements in various ways to emphasise them, while mentioning background and secondary elements as such towards the end;
- Starting with an existing prompt, either created by themselves or others, and iteratively modifying it (one user mentioned maintaining an archive of prompts for this purpose);
- A specific technique for using Midjourney involves incorporating seeds into prompts; seeds are specific inputs or initial parameters that help guide the AI model's creative direction, enabling the generation of similar and consistent designs in subsequent generations.

One user declared to **intentionally use vague descriptions in their prompts to elicit “more original inspiration”**, preferring this approach over directly “influencing the AI bot” to follow a specific direction.

Users are split on their approach to creating prompts for AI tools, with an equal number favouring a consistent method across all platforms (11 responses) and those preferring a tailored interaction for each tool (9 responses). Therefore, not all users change their approach when using different AI image generators. A quarter of respondents reported using only a single tool. However, some users demonstrate a deep understanding of how to engage with the tools, noting not only the variability of results across different platforms, but also how **identical inputs can produce different results even when submitted consecutively**. They emphasise how their prompt construction is influenced by the “language that the tools speak”, their particular “strengths in interpreting specific words”, the robots used and even the datasets on which the models are trained. The platform's interfaces significantly influence user experience, leading to distinct preferences for specific AI tools. For instance, Midjourney is favoured for its exceptional ability to produce hyper-realistic images, while DALL-E is recognized for its ability to support an interactive feedback loop within chatbot-based interfaces.

Summary

Image research seems to play an important role in the process, supporting the designer as a **source of inspiration** especially during the ideation and then the development phase. Participants tend to select inspirational images for UI design projects based on the consistency of style with the brand and the **relevance of the image to the project theme**. Images also serve a key role in **suggesting "original" design solutions for specific components**, with visual quality significantly influencing selections, but possibly for moodboard creation. Also **colour palettes** are frequently extracted from images, highlighting colour's critical role in influencing design decisions.

Surprisingly, more than half of the respondents experienced **difficulties in finding suitable inspirational images** due to a number of factors. Even in traditional online image search, **defining effective keywords** is a challenge. Designers often struggle to search for abstract concepts or specific visualisations that truly **capture the essence of their design ideas**. Furthermore, the **vast array of results** and choices **can overwhelm** users, leaving them in a perpetual search for a better alternative. Finally, to address the risk of plagiarism, users have to **avoid direct references to online projects** in their creative process.

Almost half of the survey participants have not yet incorporated AI-based image generation into their design workflow. Three-fifths of this group believed AI could assist in the ideation phase, and a comparable number recognized its potential for creating detailed presentation imagery. However, there was **scepticism regarding AI's effectiveness in defining possible solutions**.

Qualitative feedback revealed that some designers may have a **specific concept of applying AI that hasn't yet been proven fully effective**, leading them to not integrate the technology into their projects **without exploring its potential applications** at different project stages. One notable recommendation is the use of AI for fast image production, with the aim of rapid prototyping. However, the predominant perspective among these users is to see AI mainly as **a tool to enhance visual content, rather than for helping to generate design solutions**.

Among the 27 designers who utilise AI, the majority employ it for inspiration and for solution development, indicating a preference for **applying the technology earlier in the process rather than towards its completion**. However, only a few designers apply AI directly to UI screens and digital interfaces, suggesting a broader use for generating abstract inspirational images or moodboards.

In this specific user group, Midjourney and DALL·E 3 (accessed by both Bing Image Creator and ChatGPT 4) are the most popular and the favourite image generator tools. Designers selected AI tools prioritising the **aesthetic quality** of generated images, their **correspondence with the given prompts**, and the capability to produce images in high resolution, addressing key professional requirements. The capacity for **generating a wide range of images** to meet diverse design needs is also valued, with users expressing a slight preference for tools that offer conversational input methods over those requiring keyword lists. Yet, the simplicity of the tool's interface was not identified as a decisive factor. There is a lack of interest among the user group in community engagement, despite its potential to improve interactions and results.

Finally, feedback emphasises that UI design goes beyond mere aesthetic arrangement and the capabilities of image generators, underscoring the **critical role of professional insight and iterative processes in refining AI-generated outputs into final, user-centric designs.**

Most users reported independently crafting text prompts, but only about half used this approach exclusively. This suggests they **navigate AI platforms through trial and error**, blending various techniques without a definitive interaction methodology. Strategies include **seeking tips or guidelines** and **consulting other AI tools** like ChatGPT.

Participants shared several suggestions for prompt creation:

- **Utilising short sentences or keywords**
- **Emphasising features they want to highlight in the output by repeating key elements and clearly labelling secondary elements as such**
- **Beginning with a pre-existing prompt and making iterative adjustments**
- **Integrating seeds into prompts to enable the generation of similar and consistent designs in subsequent generations**
- **Intentionally use vague descriptions to elicit “original inspiration” from the AI**

While not all users modify their strategies across different AI image generators, some adeptly adjust their prompt creation to align with the unique "language" and interpretive strengths of each tool, as well as their varying interfaces. This tailored approach leads them to have preferred use cases for specific platforms.

Ethnographic research findings

Online ethnographic methods were conducted to gain qualitative insights into the specific question: what are the concepts and perceptions surrounding the use of Artificial Intelligence in User Interface Design?

The initial idea was to differentiate opinions on the use of simple image generators and editors (or UI generators, as uizard is defined) for the production of graphical layouts. From the data collected, it became clear that the latter are not really being used for the purpose they have been conceptualised, i.e. to finalise the UI with high-definition graphical components arranged according to a proposed layout. Generating suggestions from textual prompts helps in the initial idea generation phase, in overcoming the fear of the blank page and in getting started with the first embryonic graphic elements.

At the moment, therefore, there does not seem to be much difference in usage between the two types of tools, which are employed for their ability to provide creative inputs. For this reason, opinions and main themes were collected and compared, and no massive differences were found.

It should not be forgotten that the action of rapidly generating prototypes and editable designs lacks, at least for the time being, an evaluation of the user experience. Therefore, having a few screen proposals should not be confused with the actual design of the product in terms of navigation flow and information architecture.

Thematic Analysis

Opinions and **patterns of ideas** towards the topic were extracted from user comments, analysed and systematically categorised. Remarkably, four of the six themes identified had previously been noted in both the literature review and the survey results, so that these **user perceptions were confirmed by the findings of the ethnographic study**. On the other hand, **two themes emerged as entirely new**: a perceived appropriation of the creative aspect of the process, and a perspective that questions the overall utility of the technology.

The most polarised views have more in common than the moderate ones, with recognising AI as a breakthrough and disruptive technology. However, they diverge in their predictions for the future impact on designers and, more broadly, on different professions. Those in the second category often compare AI to past inventions such as photography or the introduction of software, which historically shifted practices from manual to digital methods. They appear to welcome the incorporation of new technologies into their workflow, viewing them, much like historical advancements, as a natural and progressive evolution and transformation in human labour.

The following table summarises the findings, highlighting expected and unexpected results. **Concepts** that are **unique to a particular tool category** (image generators or UI editors/prototyping tools) are marked with a visual reference.

Hints

Patterns

AI is seen as a game-changer that will empower designers who use it over those who don't

AI is an inspirational and imaginative tool for generating new ideas, giving its own interpretation of the prompt

<<[...] The most critical part of a web design is the inspiration itself. Getting a genuine inspiration of a unique design is more than enough to get started with.>>; <<[...] here's the crazy thing: it kind of inferred the type of features that you might find inside of the application [...]>>

The role of the designer may evolve as responsibilities and parts of the work are delegated to AI

<<[...] it somehow also holds a sort of beauty to in the future I become even more of a creative director using tools to create the vision (or let them create) [...]>>; <<[...] we can start offloading design debt [...]>>

AI's ability to quickly analyse data is a significant time saver, freeing up time for other tasks

<<I usually scroll through unsplash for hours for getting ideas [...]>>; <<[...] now I get to work on bigger problems, I get to go deeper. So, instead of minimum viability, I can work on maximum viability[...]>>

AI is seen as a simple tool and a technological improvement, providing automation within the process

Incorporating AI-generated results into a practical design requires diverse expertise

<<Getting the graphics usable by a typical front end developer takes a Good amount of work.>>; <<[...] the AI is also using references that do NOT apply good design techniques, and require experts to make discernments instead.>>; <<[...] you have to have a vision, a clearly defined problem in mind before you can really use these tools to your advantage.>>

AI suggestions focus on UI and not UX, lack of consideration of various factors

<<This does aesthetics extremely well. It can't do usability or UX [...]>>; <<[...] a good designer does something more than just moving beautiful pixels around. It's always about overcoming restrictions and limitations that come from the nature of the business and gets dictated by its goals. [...]>>

AI is able to generate solutions for simple interfaces, not complex interactions

<<The resulting "websites" seem to be only suitable for basic content [...]>>

AI application is seen as pointless and ineffective

AI doesn't add value as an alternative to current methods

<<I don't see a difference between this, and the thousands of free and paid templates out there [...]>>; <<if you cannot export the assets of these landing pages designs, it is not that useful>>

Results lack authenticity and don't communicate value

<<These designs are visually appealing but substance-poor and generic. [...]>>; <<everything pretty much looks the same, like it was generated, it's missing some of the finer details and creativity from a human touch>>

AI tools are only valid for creating hero or stock images

<<Am I missing something? So you used AI just to create your hero image right?>>

AI application requires regulation to protect intellectual property

The models are trained on publicly accessible content

<<[...] I would be furious to see images online heavily inspired by my work. [...] Some artists have a really strong style that can be easily recognized. [...]>>

No permission is requested, nor is credit attributed to the original source material

<<We recognize a difference between inspiration and taking a design and just switching a few things around without acknowledging the source.>>

AI production is seen as unfamiliar and intrusive

The AI production is easily identifiable and repulsive

<<[...] I always can tell the difference between human and AI art. All the AI production is the same and it has some Dark Alien feel to it! [...]>>

AI takes and appropriates creative human work

<<So where is your creativity if AI is doing all the work for you...?>>; <<This is how AI takes creativity and brainstorming from us as developers [...]>>

AI's future impact on the role of the designer, and more broadly on human work, is disruptive

The incredible speed of technological development could lead to the substitution of human labour

<<so AI will replace UI/UX designers?>>

Automation could reduce production costs, leading to an increase in demand for production levels

<<[...] you spent extremely small amount of time on it, reducing costs for conceptualising and getting the asset. The question is if it just makes demands bigger.>>

People with management responsibilities may not see the value in the work of a designer employee

<<[...] Designers aren't the ones deciding if AI is good enough to replace our jobs. It's the employer who think AI work is good enough. [...]>>

AI has the ability to summarise vast amounts of data in an unbiased way



AI outputs are not final designs and leave room for the designer's creativity and interpretation

Decision-making processes are supported and accelerated thanks to prototyping and visualisation of ideas

EXPECTED

<<[...] it (AI) summarizes some (works) into a collage - something you'd normally do yourself as moodboards. [...]>>; <<[...] this is what AI is doing right now by parsing and collaging unbiased trends [...] .>>

<<[...] The actual UI patterns are quite useless from AI, and that's the entire value of the source material.>>

<<It's very normal for us to think of like five, ten ideas at once and you're just not fast enough and when you start working on one thing you start getting an idea for another thing...in this case it's so exciting to have an AI companion that sort of help you move faster>>

Automation and technology improvements are a natural progression and must be embraced

There is a lot of room for improving these tools to be more helpful in the design process

As AI relies on existing data for training, there is a need to provide new ideas for future model development

EXPECTED

<<[...]The best way to future-proof your career is to embrace change early rather than fight it. [...] New industries and jobs will continue to emerge.>>

<<I would suggest for this app to feed it not only prompts, but data, personas, IA, etc then it would be cool. [...]>>; <<Great point. For me, the way to address this is to have AI design the mobile view first, then do the manual work to adapt the design to desktop. This is a far simpler exercise.>>

<<[...] what it's creating today it looks like things that are already existing today and if the training materials never get updated with new innovative trends new aesthetics, it never changes... like everything's gonna look the same[...], we still need talented creative people out there making new stuff to train the new AI to maintain or to keep pushing forward [...]>>

In the repetitive results there is no creative contribution



<<It's not really "custom", it seems that they have a few templates that are being randomized [...]>>

UNEXPECTED

EXPECTED

UNEXPECTED

EXPECTED

UI editor

Image generator

Summary

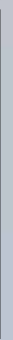
AI has the potential to **transform the role of designers**, according to some users, thanks to its **unbiased summarising capabilities** and **decision-making support**. Its outputs offer a **distinct interpretation of prompts**, such as deducing features of a described application, while still allowing for creative input. Currently, **image generators and UI editors have a very similar application** in their ability to provide insights and suggestions in the early stages of idea generation.

Some see **AI integration** not as a breakthrough innovation, but **as a process that still requires team skills and knowledge**, as well as **a defined project development strategy that the technology itself cannot provide**. The results should not be considered final, due to the lack of UX evaluations, business considerations and solutions for complex interactions. Nevertheless, technological progress is seen as a natural progression of human work.

AI is occasionally viewed as superfluous and ineffective compared to current methods, offering no real creative value: UI editor **outcomes are criticised for being inauthentic, repetitive, and generic**. At most, image generators are capable of producing hero or stock images. In the most extreme case, generated content is perceived as distinctly recognisable and alien, largely due to concerns that technology is usurping the creative domain traditionally reserved for humans.

Many professionals are worried about copyright infringement, pointing to the **lack of adequate safeguards for publicly available content on the web used to train models**. AI-generated output does not give credit to the original source material.

Concerns remain that automation will not only increase production demands, but also replace human designers entirely. This **fear** arises not only **from the capabilities of the technology**, but also from a possible **diminishing recognition of the value of the human touch** among those in management roles.



Triangulation

results

Define the role in the design process

The primary objective of this research was to explore the potential role and value that AI tools could add to the design process. Existing literature does not provide clear guidance on how designers should integrate AI-generated content into their workflows. A survey method was chosen for its unique contribution to gathering insights beyond those found in the literature, focusing specifically on **how professionals select inspirational images and tools for image generation based on their needs**. Often, AI is seen as a **means to produce visual content**, such as hero or stock images, **rather than as a resource to facilitate design solutions**. Furthermore, the reluctance to adopt AI tools for applications that are not fully effective contributes to their exclusion from the design process.

While some studies view AI generation as an alternative to **traditional image search**, others argue for a **complementary** relationship, recognising the distinct benefits of each approach. Both are valued for inspiration in the ideation and development stages, with AI tools particularly noted for their ability to enhance creativity across disciplines. A key area of interest lies in the unique obstacles and opportunities presented by each method: **creating effective prompts for image generation and identifying relevant keywords for search queries have both emerged as notable challenges**. The use of AI-generated images as “end products” raises **ethical concerns about the source of the training data** (often public content), as well as the **potential for plagiarism with direct online references**. Therefore, there are no definitive disadvantages of AI generation when compared to conventional online image searches in the literature. Strengths, on the other hand, are better able to define the benefits of potential applications. The power of AI lies in its ability to **additively combine elements from detailed prompts in a constrained generative process**, as opposed to the array of filtered and specific results that online search queries produce. This traditional method is more suitable for discovering a wide range of precise and well-defined examples of specific graphical components. AI excels at **representing abstract concepts** through less defined images, thanks to its ability to blend styles and evoke emotional nuance. **The inherent vagueness of textual prompts offers a range of unpredictable and varied outcomes, encouraging divergent thinking**. This element of uncertainty is particularly beneficial, as it may inspire creativity more than precise, polished suggestions.

The literature distinguishes two categories of AI-supported tools in the design process: **image generators**, which, although not exclusively for UI design,

incorporate UI images into their training datasets, and **UI editors**, specialised platforms for creating user interfaces that have adopted AI functionalities. In particular, the latter was designed to automate the manual and time-consuming creative process by suggesting and refining layouts for faster delivery. However, the digital ethnography revealed that **both types now play a similar role, particularly in the early stages of ideation.**

This similarity arises because **their output should not be seen as final, lacking UX evaluation, business insight and complex interaction solutions.** In the realm of UI editors, AI can suggest editable layouts, but these often turn out to be repetitive, generic and lacking in stylistic cohesion, requiring further designer refinement and touch. Images suggested by generators need to be converted into modifiable graphical components across platforms, with a selection process to filter out unwanted clutter for the final interface. Integrating AI-generated ideas into practical designs requires a wide range of skills and alignment with project objectives and organisational dynamics.

The analysis from all three research methods underscores **the essential role of the designer's informed evaluation, a function that AI cannot yet replicate.** This assessment goes beyond basic aesthetic considerations and emphasises the transformation of AI-generated ideas into polished, user-focused designs. In this human-machine collaboration, computational tools are envisioned to **augment human creativity by providing novel inputs**, such as inferring features for a described application. As AI takes on more responsibility and workload, the role of the designer may shift, taking advantage of its **impartial analysis and summarisation capabilities.** The potential for rapid iteration and prototyping enhances the decision-making process, allowing designers to take their solutions **from “minimum viable” to “maximum valuable”.** While automation can lower production costs, it can also raise output requirements and reduce the perceived value of a designer's work, particularly in the eyes of management. The future impact depends on the ability of professionals to **adapt their skills to new technologies** and their intended applications.

Facilitate interaction with the tools

From all three research methods, it's clear that users primarily engage with AI tools through a **trial and error approach, seeking guidance or consulting other AI resources such as ChatGPT for prompt generation**. This preference for conversational input methods is likely due to this exploratory nature of interaction. A key insight from the research is the deliberate use of **intentionally vague descriptions** by some users to **spark "original inspiration"** from the AI.

From the results, specific recommendations are derived for future improved interaction. The aim of this guide is to offer suggestions **without restricting creativity and the adaptability of text-to-image AI technologies**. Recommendations include:

- **Use concise sentences or keywords.**
- **Emphasise key features through repetition and clear labelling of secondary elements.**
- **Incorporate "prompt specifiers" to produce output that is both precise and conceptually aligned, allowing more than one word to be integrated for each. The synthesised UI design taxonomy includes:**
 - **Style and Mood:** attributes that add context and can express the image's abstract qualities, like emotional tone or stylistic inclination.
 - **Colour:** global colour schemes that usually precede the main subject, specifying particular palettes or broader categories (e.g., colourful, monochrome).
 - **Subject:** the core elements and focus of the image, possibly specifying device types (e.g., desktop, mobile app), page or interface types (e.g., e-commerce, landing page), and main themes (e.g., fashion, food delivery).
 - **Medium:** the requirement of a UI indicates a photographic representation of digital graphic and visual elements; these may extend to different artistic media (e.g. illustrations, drawings).
 - **Concrete Details:** specific features or elements desired in the UI.
 - **Influence:** the artistic style is influenced by references to artworks, artists (as well as celebrities from related fields, such as music and fashion), art repositories, art genres, movements, and platforms (e.g., Pinterest, Dribbble), often listed sequentially.
 - **Composition:** general layout considerations (e.g., multiple digital screens, plain backgrounds), aspect ratios (e.g. ar 16:9, ar 9:16), perspectives (e.g., flat view), and negative prompt (e.g., no monitor).
 - **Quality Boosters:** keywords that can enhance image quality (e.g., "beautiful website," "high resolution").
- **Use synonyms for subjects or varying sentence constructions for achieving distinct styles.**
- **Start with an existing prompt for iterative refinement.**
- **Integrate seeds or constants (like style cues) and variables (specific subjects or details) to guide the generation of similar, cohesive subsequent designs.**

Propose new evaluation criteria

Currently, there are no consistent guidelines for the qualitative evaluation of AI-generated models. As AI becomes increasingly **integrated into the design industry**, there's a critical **need for a comprehensive set of evaluation criteria that are specific to the domain and professional requirements**. In addition, the establishment of new criteria could deepen any understanding of the limitations that exist in generating the specific theme.

One of the major contributions of this research is the comparison and selection of metrics from various literature studies with the image and tool selection criteria identified through the active participation of professionals.

The primary metrics traditionally used in evaluations focus on the realism of the generated image (image fidelity) and how well the image aligns or semantically correlates with its description. The latter is essential for **sourcing inspirational images that resonate with the project theme and adhere to the brand's style**, elements that are often outlined in the prompt, making this a crucial criterion. While consistency in the results is desirable, **too much similarity can lead to design fixation**, highlighting the importance of evaluating the diversity within the results generated. The potential of AI to suggest original designs, whether in layout, graphic elements or colour schemes, plays a significant role in the selection process. The images generated should not only inspire, but also offer **originality**, as recognised by the HEIM benchmark, and **novelty to avoid copyright issues**. On the other hand, the actual realism of an image is not considered a strict necessity. Inspiration can also come from more **abstract or vaguely defined images**, which **can evoke and convey moods better than explicit solutions**. The resolution, visual quality and **overall aesthetics** of an image have a significant impact on designers' choices and perceptions.

Thus, the newly proposed metrics for evaluation include:

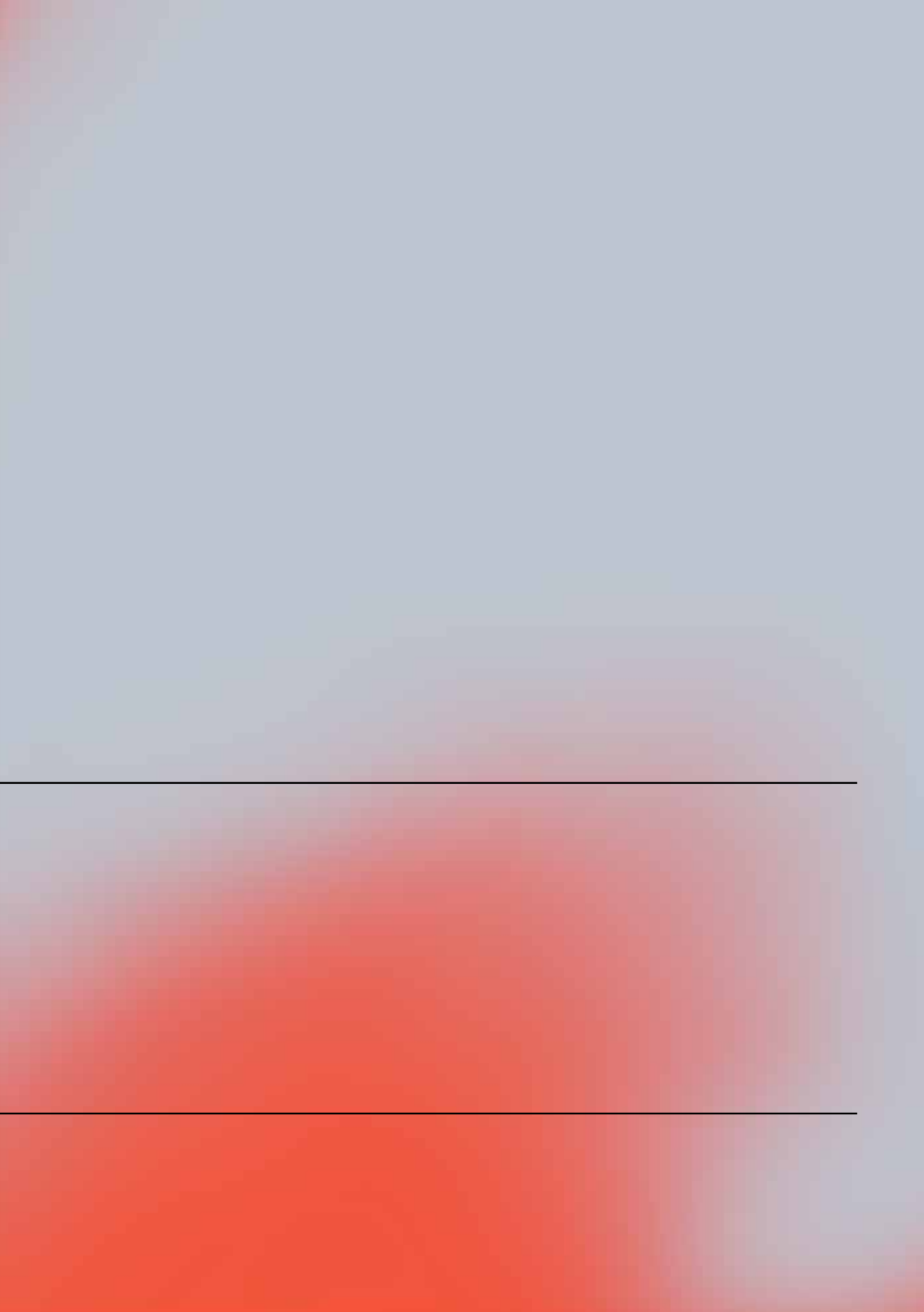
- **Alignment with the prompt, ensuring thematic relevance and stylistic consistency**
- **Diversity of the generated outcomes**
- **Contribution to inspiration**
- **Originality, uniqueness, and novelty**
- **Resolution and visual quality**
- **Overall aesthetics**

CONCLUSIONS

3.1 *Process*

3.2 *Results*

Critical analysis
Limitations
Future steps



Process

The research was conducted using a multi-method approach. Due to the rapid evolving nature of the topic, it was essential to compare information from various sources. Initial steps in academic research have been made, particularly in different design areas. However, the literature review identified gaps, such as the absence of meaningful and standardized evaluation criteria that address designers' specific needs. Although academic research advances theoretical knowledge and assumptions about tool capabilities, practical validation is necessary. To address this, field professionals were surveyed to assess the value of traditional image search and the potential of generating visual cues, both images and layouts, as complementary tools. Additionally, digital ethnography was employed to explore patterns in users' ideas, judgments, and perceptions, verifying the practical applications and capabilities of these tools in real-world professional settings.

Results

Critical analysis

This research has consistently argued that integrating generative AI into the professional design domain is crucial for new developments, emphasizing the importance of involving professionals in collecting requirements. In the absence of guidelines and relevant metrics, there is a critical need to provide practical validation of current applications to the theorised and assumed capabilities in academia. A significant contribution of this study is the comparison of literature data with the opinions of surveyed professionals, providing a comprehensive overview of the current state of the art. Furthermore, suggested metrics based on user needs can deepen the understanding of the limitations in generating the specific theme.

The literature distinguishes between two types of tools: UI editors and generic image generators. UI editors create editable layouts with inferred features, often resulting in repetitive and generic outputs lacking stylistic cohesion. Generic image generators produce UI pictures that need to be converted into modifiable graphical components, requiring a selection process to eliminate unwanted elements for the final interface. Beyond perceptions of AI tools, **both their strengths and inherent limitations support the use as a source of inspiration**, particularly in the early stages of ideation.

AI's ability to provide **rapid iteration and prototyping** supports decision-making processes. It encourages **divergent thinking** and enhances creativity by offering a complementary role to traditional image searches. A well-known challenge is creating effective prompts for generation, where the vagueness of textual descriptions results in **varied and unpredictable outcomes**. This unpredictability can inspire creativity more than precise, polished suggestions. Since the potential for design fixation is a concern when outputs are too similar, it is important to consider the metric of diversity. However, it is also important to ensure that the alignment with the prompt is evaluated, in order to guarantee thematic relevance and stylistic consistency.

Additionally, AI excels at **representing abstract concepts** and emotional nuances, combining elements, and **blending styles**, which can **inspire original designs** in layout, graphic elements, or color schemes. The realism of an image is therefore less important than its ability to inspire, although resolution and overall aesthetics significantly impact designers' choices and perceptions. This research gathered data on current strategies that support interactions with AI tools. While “prompt specifiers” taxonomies and using ChatGPT for prompt generation are valid

options, an interesting outcome is the deliberate use of intentionally vague descriptions to spark “original inspiration” from AI. These guides and supports must not restrict the adaptability that text-to-image technology offers.

Moreover, the intellectual origins of publicly available material used to train models are often unknown and not referenced. The designer’s evaluation and interpretation of AI outputs could mitigate **plagiarism concerns associated with direct inspiration** from online resources. Novelty and originality metrics can help avoid copyright issues.

In conclusion, despite the distinction between tool types, both **lack UX evaluation, business insight, and complex interaction solutions**. Integrating AI-generated ideas into practical designs requires diverse skills and alignment with project objectives and organizational dynamics. Therefore, the designer’s informed evaluation in an **iterative and non-linear process** remains essential for creating user-centered designs and cannot yet be replicated. The role of the designer is shifting towards a “**machine creation with designer evaluation**” model. This evolution could transform solution development from “minimum viable” to “maximum valuable” creation.

Limitations

This research has **limitations in both methodology and scope**. The extensive number of variables considered means this study offers only a modest contribution toward a comprehensive understanding of AI-generated visual cues in design of digital interfaces.

Methodologically, the small sample size of professionals surveyed **limits the ability to draw quantitative conclusions**. Additionally, the limited qualitative data from previous studies and the **lack of standardized evaluation criteria** make it challenging the comparison and analysis of the knowledge produced. The final sample of articles reviewed is relatively small, and the **vast amount of potentially available data** from non-scientific sources, such as blogs, social platforms, and media, is immense and challenging to encompass.

General obstacles also affect the conclusions drawn from this research, necessitating a broader evaluation of AI applications. A significant issue is the nature of AI generation, with potential plagiarism concerns arising from the data used for model training. A comprehensive evaluation and **regulatory framework** are necessary to guarantee the originality of AI-generated content and **to prevent copyright infringement**, including when these generations are used for inspiration in the early stages of design (one of the suggestions supported by this research). Additionally, the interpretation of AI-generated results by different professionals needs further exploration. The **widespread availability of these tools necessitates a rigorous and discerning approach to their application in the professional world**. One concern is that the visual suggestions generated by these tools may lead to the standardisation of future creative production in areas such as UI design.

Future steps

The identified obstacles currently limit the ability to draw broader conclusions at this stage. However, they offer a valuable foundation for future exploration and reflection.

With the rapid pace of technological development affecting all areas of life, the design field cannot ignore the significant benefits offered by artificial intelligence. However, it is crucial to explore how this technology can be effectively applied in the design process and how it can change the role of the professional. The future impact will depend on the **ability of designers to adapt their skills to new technologies** and their intended applications.

Future research should **involve designers in the direct comparison and testing of these tools using a standardised process** that requires prompts to be developed from the same brief. This could include a list of recommendations for prompt engineering based on the findings of this study and evaluation of the results using suggested metrics. Such research would **inform the development of new functionality and user interactions**, ensuring that they better meet user needs.

APPENDIX

4.1 *List of reference*

4.2 *List of figures*

List of figures

List of tables

4.3 Appendix

Appendix 01
Appendix 02

List of reference

A

- aaoa Streams (2024) Exploring Galileo AI: An AI-Powered Design Tool. 11 February. Available at: <https://www.youtube.com/watch?v=vSA3-Xw2xGs> (Accessed: 3 March 2024).
- Abdallah, Y.K., Estévez, A.T. (2023). 'Biomaterials Research-Driven Design Visualized by AI Text-Prompt-Generated Images', *Designs*, 7(2). doi:10.3390/designs7020048
- Alban, K. (2023) 2 A.I. Design tools to design faster. 12 February. Available at: <https://www.youtube.com/watch?v=kVKg5DG5Rpl&t=146s> (Accessed: 3 March 2024).

B

- Babich, N. (2023) 'UI Design using Midjourney', *Medium*, 27th February. Available at: <https://uxplanet.org/ui-design-using-midjourney-713e8d1b6a6b> (Accessed: 11 January 2024).

C

- Cai, A., Rick, S.R., Heyman, J., Zhang, Y., Filipowicz, A., Hong, M.K., Klenk, M., Malone, T. (2023). 'DesignAID: Using Generative AI and Semantic Diversity for Design Inspiration', *Proceedings of the ACM Collective Intelligence Conference*, CI 2023, pp. 1-11. doi:10.1145/3582269.3615596
- Çiçek, S., Turhan, G.D., Özkar, M. (2023). 'Reconsidering Design Pedagogy through Diffusion Models', *Proceedings of the International Conference on Education and Research in Computer Aided Architectural Design in Europe*, 1, pp. 31-40. doi:10.52842/conf.ecaade.2023.1.031
- Codex Community (2022a) AI Art Web Design is Amazing - Are AI Website UI The Future?. 21 December. Available at: <https://www.youtube.com/watch?v=fFMSmmW1LIY> (Accessed: 3 March 2024).
- (2022b) How to use AI Art and ChatGPT to Create a Insane Web Designs. 27 December. Available at: <https://www.youtube.com/watch?v=8I3N-TE4cn5s&t=374s> (Accessed: 3 March 2024).
- (2023a) How to use AI Art to Redesign Websites with MidJourney. 24 January. Available at: <https://www.youtube.com/watch?v=stFORMbOQS8&t=239s> (Accessed: 3 March 2024).
- (2023b) How MidJourney 5.1 brings new changes to Web Designs AI Art. 20 May. Available at: <https://www.youtube.com/watch?v=0igwt0djJEI> (Accessed: 3 March 2024).
- (2023c) New AI Art Web Design with MidJourney 5 (v4 versus v5). 24 March.

- Available at: <https://www.youtube.com/watch?v=gVA3FCqy7lc> (Accessed: 3 March 2024).
- (2023d) Reacting to Beautiful AI Art Web Designs + How to build them. 10 January. Available at: <https://www.youtube.com/watch?v=qi1lrnvc-EA&t=1013s> (Accessed: 3 March 2024).
- (2023e) 10 Web Design Tricks using Midjourney. 7 July. Available at: <https://www.youtube.com/watch?v=mEKEyKw00LM> (Accessed: 3 March 2024).
- Cornell University Library (2021) Tutorial: Evaluating Information: Gray Literature. Available at: <https://guides.library.cornell.edu/c.php?g=543699&p=3725403> (Accessed: 28 January 2024).

D

- DesignCourse (2023a) I asked Framer AI to Build Me 5 Websites. 16 June. Available at: <https://www.youtube.com/watch?v=hn1E3X4LATM> (Accessed: 3 March 2024).
- (2023b) Midjourney Version 5 for Web Design - The Results. 16 March. Available at: https://www.youtube.com/watch?v=Z_f9Rn69PdM (Accessed: 3 March 2024).
- DesignerUp (2023a) Uizard's Autodesigner is Here! But Has it Gotten Better or Worse?!. 14 June. Available at: <https://www.youtube.com/watch?v=VnezicJbf-PM> (Accessed: 3 March 2024).
- (2023b) Watch AI Turn a Single Text Prompt into Stunning UI Designs in SECONDS! 🤖 | Uizard Autodesigner. 3 April. Available at: <https://www.youtube.com/watch?v=4ilOvRORmTl> (Accessed: 3 March 2024).

E

- Edwards, C. (2023) Using AI for UI Design is INSANE!. 17 January. Available at: <https://www.youtube.com/watch?v=toivr57FjTw&t=12s> (Accessed: 3 March 2024).

F

- Flux Academy (2023) Framer Ai is INSANE!. 15 June. Available at: <https://www.youtube.com/watch?v=QKUnUrsDzkO> (Accessed: 3 March 2024).
- Frolov, S., Hinz, T., Raue, F., Hees, J., Dengel, A. (2021). 'Adversarial text-to-image synthesis: A review', *Neural Networks*, 144, pp. 187-209, doi:10.1016/j.neunet.2021.07.019.

G

- Goel, V. (2023) 'Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding', *Medium*, 17th November. Available at: <https://medium.com/@viditgoel3/photorealistic-text-to-image-diffusion-models-with-deep-language-understanding-5d7fbb4860ec> (Accessed: 12 January 2024).
- Gross, N. (2023) Will Galileo AI Replace Me as a Designer?. 10 February. Available at: <https://www.youtube.com/watch?v=zzYDevictBA> (Accessed: 3 March 2024).

H

- Hanafy, N.O. (2023). 'Artificial intelligence's effects on design process creativity: "A

study on used A.I. Text-to-Image in architecture”, *Journal of Building Engineering*, 80. doi:10.1016/j.jobee.2023.107999

Hegemann, L., Jiang, Y., Shin, J.G., Liao, Y.-C., Laine, M., Oulasvirta, A. (2023). ‘Computational Assistance for User Interface Design: Smarter Generation and Evaluation of Design Ideas’, *Conference on Human Factors in Computing Systems - Proceedings*. doi:10.1145/3544549.3583960

Hughes R.T.; Zhu L.; Bednarz T. (2021). ‘Generative Adversarial Networks–Enabled Human–Artificial Intelligence Collaborative Applications for Creative and Design Industries: A Systematic Review of Current Approaches and Trends’, *Frontiers in Artificial Intelligence*, 4. doi:10.3389/frai.2021.604234

K

Karahan, H.G., Aktaş, B., Bingöl, C.K. (2023). ‘Use of Language to Generate Architectural Scenery with AI-Powered Tools’, *Communications in Computer and Information Science*, 1819 CCIS, pp. 83-96. doi:10.1007/978-3-031-37189-9_6

Kenthapadi, K., Lakkaraju, H., Rajani, N. (2023). ‘Generative AI meets Responsible AI: Practical Challenges and Opportunities’, *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 5805-6. doi:10.1145/3580305.3599557

Koh, I. (2023). ‘Architectural and Social ‘Word Forms’ Slippage with Deep Learning’, *Communications in Computer and Information Science*, 1819 CCIS, pp. 112-25. doi:10.1007/978-3-031-37189-9_8

L

Lee, Y-H., Chiu, C.-Y. (2023). ‘The Impact of AI Text-to-Image Generator on Product Styling Design’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 14015 LNCS, pp. 502-15. doi:10.1007/978-3-031-35132-7_38

Lee, T., Yasunaga, M., Meng, C., Mai, Y., Park, J.S., Gupta, A., Zhang, Y., Narayanan, D., Teufel, H.B., Bellagente, M., Kang, M., Park, T., Leskovec, J., Zhu, J., Fei-Fei, L., Wu, J., Ermon, S., Liang, P. (2023). ‘Holistic Evaluation of Text-to-Image Models’, *Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track*.

Li, J., Yang, J., Hertzmann, A., Zhang, J., Xu, T. (2019). ‘Layoutgan: Generating graphic layouts with wireframe discriminators’, *7th International Conference on Learning Representations*, ICLR 2019.

M

Manovich, L. (2022) ‘Chapter 4 - AI and Myths of Creativity’, *Artificial Aesthetics: Generative AI, Art, and Visual Media*. Available at: <http://manovich.net/index.php/projects/artificial-aesthetics-book> (Downloaded: 18 January 2024).

Mathiesen, S.O., Canossa, A. (2023). ‘If You Can’t Beat Them, Join Them: How Text-to-Image Tools Can Be Leveraged in the 3D Modelling Process’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 14059 LNCS, pp. 162-181. doi:10.1007/978-3-031-48057-7_11

Miranda, L.J. (2021) *The Illustrated VQGAN*. Available at: <https://lvmiranda921.github.io/notebook/2021/08/08/clip-vqgan/> (Accessed: 11 January 2024).

Mountain, J. (2023) Using DALL-E 3 To Design My Software UI: Does It Work?. 31 October. Available at: https://www.youtube.com/watch?v=KD1VO_Br68E

(Accessed: 3 March 2024).

Mozaffari, M.A., Zhang, X., Cheng, J., Guo, J.L.C. (2022). 'GANSpiration: Balancing Targeted and Serendipitous Inspiration in User Interface Design with Style-Based Generative Adversarial Network', *Conference on Human Factors in Computing Systems - Proceedings*. doi:10.1145/3491102.3517511

N

Nguyen, T.T., Vu, P.M., Pham, H.V., Nguyen, T.T. (2018). 'Deep learning UI design patterns of mobile apps', *Proceedings - International Conference on Software Engineering*, pp. 65-8. doi:10.1145/3183399.3183422

O

OpenAI (2021) *DALL-E: Creating images from text*. Available at: <https://openai.com/research/dall-e> (Accessed: 11 January 2024).

P

Pieters, R., Wingers, S. (2016) 'Creative AI: On the Democratisation & Escalation of Creativity', *Medium*, 7th March. Available at: <https://medium.com/@creativeai/creativeai-9d4b2346faf3> (Accessed: 13 January 2024).

S

Saharia, C., Chan, W., Saxena, S., Li L., Whang, J., Denton, E., Ghasemipour, S.K.S., Ayan, B.K., Mahdavi, S.S., Gontijo-Lopes, R., Salimans, T., Ho, J., Fleet, D.J., Norouzi, M. (2022). 'Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding', *Advances in Neural Information Processing Systems*, 35.

Sanchez, T. (2023). 'Examining the Text-to-Image Community of Practice: Why and How do People Prompt Generative AIs?', *ACM International Conference Proceeding Series*, pp. 43-61. doi: 10.1145/3591196.3593051

Showalter, J. (2023) Is this the end of UI/UX Designers? | Auto Designer by uizard. 19 June. Available at: <https://www.youtube.com/watch?v=PjcP0eXyQQ0> (Accessed: 3 March 2024).

Siddiqui, N. (2023) *Comparative Study of Generative Models for Text-to-Image Generation*. Thesis. University of Windsor.

Stigsen, M.B., Moisi, A., Rasoulzadeh, S., Schinegger, K., Rutzinger, S. (2023). 'AI Diffusion as Design Vocabulary: Investigating the use of AI image generation in early architectural design and education', *Proceedings of the International Conference on Education and Research in Computer Aided Architectural Design in Europe*, 2, pp. 587-96.

U

Urietti, L. (2023) *Towards algorithmic intimacy. Machine learning generativo applicato al branding*. Thesis (Master). Politecnico di Milano.

V

Verheijden, M.P., Funk, M. (2023). 'Collaborative Diffusion: Boosting Designerly Co-Creation with Generative AI', *Conference on Human Factors in Computing*

W

Wilson, D. (2022) I Generated 80 INSANE AI Web Designs in 1 Hour. (SHOCKING RESULTS!). 20 December. Available at: https://www.youtube.com/watch?v=kuYw_VYrDSs (Accessed: 3 March 2024).

Wu, F., Hsiao, S-W., Lu, P. (2024). 'An AIGC-empowered methodology to product color matching design', *Displays*, 81. doi:10.1016/j.displa.2023.102623

Y

Yang, H., Xue, C., Yang, X., Yang, H. (2021). 'Icon generation based on generative adversarial networks', *Applied Sciences (Switzerland)*, 11(17). doi:10.3390/app11177890

List of figures

List of figures

| | | |
|------------------|--|----------------|
| <i>Figure 01</i> | Methods and objectives | <i>p.17</i> |
| <i>Figure 02</i> | Prompt specifier taxonomies | <i>p.35</i> |
| <i>Figure 03</i> | Comparing criteria used in different studies | <i>p.38-39</i> |
| <i>Figure 04</i> | Insights into the use and application of online image research | <i>p.45</i> |
| <i>Figure 05</i> | Have you already integrated (or tried to integrate) AI-based image generation tools into your design workflow? | <i>p.48</i> |
| <i>Figure 06</i> | At which stage of your design process have you utilized AI-based image generation? | <i>p.48</i> |
| <i>Figure 07</i> | At what stage do you envision AI tools could be beneficial? | <i>p.49</i> |

List of tables

| | | |
|-----------------|---|----------------|
| <i>Table 01</i> | Reference list for image generator videos | <i>p.23</i> |
| <i>Table 02</i> | Reference list for UI editor videos | <i>p.24</i> |
| <i>Table 03</i> | Findings from digital ethnography | <i>p.56-57</i> |

Appendix

Appendix 01

The following is a list of the channels through which the survey was distributed:

| | |
|-------------------------|---|
| Discord server | AI forum for India |
| Facebook group | UX & UI DESIGNER COMMUNITY |
| Linkedin group | Creativity in Graphic Design, Logo Design, Web, UI/UX Design, Advertising & Art Direction |
| Reddit community | r/AIArtwork |

Appendix 02

The questionnaire was named "*Application of AI-generated images in UI design_Master's thesis research*". An introduction was included at the beginning to introduce the topic and context of the research:

"Thank you for landing here!

I'm Virginia, a student approaching the conclusion of my Master's program in "Digital and Interaction Design" at Politecnico di Milan.

*This survey is **aimed at UX/UI designers**. I will first ask you about your design process and how you use **image research**: its role, importance and selection criteria. This is followed by questions about the use of AI.*

*The core of my study lies in the creation of **text prompts** for **AI image generation**, specifically applied to the **UI Design** process. All questions therefore relate to **text-to-image generation only**, and not to other types of AI application.*

*I sincerely value your contribution to this study. Please be assured that **your responses will remain anonymous** and will only be used for the purposes of this research."*

The following is the structure of the survey and the detailed questions for each of the sections.

Section 1: General introduction

- **What kind of design process do you generally follow?***
 - User-Centred Design (UCD)/Human-Centered Design (HCD)
 - Design Thinking
 - Agile Design
 - Lean UX (think, make, check)
 - Goal-Directed Design
 - Other...
- **At what stage of your design process do you typically conduct image research and how can it provide the most benefit? (You can select more than one option and add in "Other" your specific applications)***
 - As a source of inspiration during the ideation phase
 - Designing solutions
 - In the definition phase, specifically for crafting detailed presentation images
 - Other...
- **How integral is image research in your design process?***
 - 1 - Not important at all
 - 2
 - 3
 - 4 - Absolutely essential

- **What criteria do you use to select inspirational images for your UI design projects? (Select all that apply)***
 - Relevance to project theme or subject matter
 - Consistency with brand identity and style
 - Colour Palette
 - Originality and uniqueness of UI elements
 - Alignment with standard layout and UX rules
 - Latest trends
 - Scalability and adaptability across different devices
 - Visual quality and resolution
 - Other...
- **Have you encountered any difficulties in finding the right inspirational images that align with a project's theme or style?***
 - Yes
 - No
- **If yes, what challenges do you face during image research, and how do you address them?**
- **Have you already integrated (or tried to integrate) AI-based image generation tools into your design workflow?***
 - Yes ([go to Section 3](#))
 - No ([go to Section 2](#))

*Section 2: "For those **who haven't used AI generation** in the design process"*

- **At what stage do you envision AI tools could be beneficial? (Select all that apply)***
 - As a source of inspiration during the ideation phase
 - Designing solutions
 - In the definition phase, for developing detailed presentation images
 - Other...
- **Do you think AI image generation would be useful for creating which of the following graphic elements? (Select all that apply)***
 - UI screens and digital interfaces
 - App icons
 - Product images
 - Promotional illustrations
 - Logos
 - Mascots
 - Other...
- [Go to Section 5](#)

*Section 3: "For those **who have used AI generation** in the design process"*

- **At which stage of your design process have you utilized AI-based image generation? (Select all that apply)***
 - As a source of inspiration during the ideation phase
 - Designing solutions
 - In the definition phase, for developing detailed presentation images
 - Other...
- **Have you applied AI image generation in the production or ideation of any of the following? (Select all that apply)***

- UI screens and digital interfaces
- App icons
- Product images
- Promotional illustrations
- Logos
- Mascots
- Other...
- **How many AI generative tools have you experimented with, and which ones specifically? (Select all that apply)***
 - Midjourney
 - DALL·E 2
 - DALL·E 3
 - Bing Image Creator
 - DreamStudio (Stable Diffusion)
 - Runway
 - Adobe Firefly
 - Other...
- **Among the AI generative tools you've explored, do you have a favourite one(s)?***
 - Yes
 - No
- **If so, can you specify which one(s)?**
- **If so, what factors contribute to your preference?***
 - Aesthetic result of the generated images
 - Capability to create a diverse range of images, accommodating various design needs
 - Consistency and alignment of the result with the prompt provided
 - Easy prompt creation using natural language and conversational descriptions
 - Easy prompt construction by comma-separated list of values/ keywords (nouns and adjectives)
 - Free use
 - Affordability for users with budget constraints
 - Intuitiveness and ease of use in the tool's interface
 - Quick processing and generation times for images
 - Ability to generate high-resolution images
 - Presence of a supportive user community
 - Regular updates and improvements from the tool provider
 - Availability of tutorials, documentation, or educational resources
 - Other...
- **Go to Section 4**

Section 4: "Prompt creation and methodology"

- **What is your approach to constructing text prompts for these tools? (Select all that apply)***
 - I independently invent the prompt
 - I do research and follow tips or guides
 - I use other AI tools (like ChatGPT) to generate the prompt
 - Other...
- **Have you developed a methodology for building the text prompt?***
 - Yes

- No
- Can you describe your methodology for the creation of the text prompt?
- Is your process for constructing prompts consistent across all tools, or does it vary?*
- I have only used one tool
- It's always consistent
- It changes based on the tool
- Can you explain why and, if it changes, how it is different? (Please specify the name of the tool used in the explanation)
- [Go to Section 5](#)

Section 5: "Thanks for your participation in this survey!"

- This is a space for comments if you would like to suggest a tool that hasn't been mentioned, or add something that you didn't get a chance to specify in the previous questions.



POLITECNICO
MILANO 1863

