UNVEILING THE ACCURACY OF AI TEXT TO IMAGE TOOLS VS. HUMAN EXPERTISE IN INTERIOR SPACE DESIGN

(A year long journey tracking and evaluating AI text to image tools revolution 2023-2024)

الكشف عن دقة أدوات الذكاء الاصطناعي (النص إلى صورة) مقابل الخبرة البشرية فى تصميم الحيزات الداخلية

(رحلة طويلة لمدة عام لتتبع وتقييم ثورة أدوات النص إلى صوره ٢٠٢٣-٢٠٢٤)

Sarah Nabih Nasif

Interior design, Faculty of Arts and Design, October University for Modern Science and Arts (MSA), Egypt

ساره نبيه نصيف فرج

قسم التصميم الداخلي – كلية الفنون والتصميم – جامعة اكتوبر للعلوم الحديثة والآداب، مصر

sara.n_nasif91@yahoo.com

ABSTRACT

In an ever-evolving world, Artificial Intelligence tools are rapidly emerging as game-changer in all fields (Interior space design included) in a way that cannot be ignored any more .The time has come to start facing the truth and start evaluating those tools' accuracy from subject matter expert's point of view and find ways to utilize it the right way by exploring how to get the best results, and by this answering so many questions, can those tools replace interior designers totally?. Can they really generate very accurate (expert level) results? What is the most reliable tool that can be used as an assistant? The study focus on evaluating the free available text to image tools only (excluding popular mid journey) following the experimental, analysis and comparative method that lead to some important findings. Most important is the perfect prompt writing strategy, Microsoft Bing being the absolute best, technical reasons behind this and much more…

KEYWORDS

Artificial intelligence **;**Text to image**;** Interior design .

الملخص

في عالم دائم التطور، تظهر أدوات الذكاء الاصطناعي بسرعة فائقة لتغير قواعد اللعبة في جميع المجالات (بما في ذلك تصميم الحيزات الداخلية) بطريقة لا يمكن تجاهلها بعد الآن. لقد حان الوقت للبدء في مواجهة الحقيقة والبدء في تقييم دقة تلك الأدوات من وجهة نظر متخصصة وإيجاد طرق للاستفادة منها بالطريقة الصحيحة من خلال استكشاف كيفية الحصول على أفضل النتائج، ومن خلال الإجابة على العديد من الأسئلة أهمها، هل يمكن لهذه الأدوات أن تحل محل مصممي العمارة الداخلية تمامًا؟. هل يمكنهم حقًا تحقيق نتائج دقيقة للغاية (مستوى متخصص)؟ ما هي الأداة الأكثر موثوقية التي يمكن استخدامها كمساعد بمراحل التصميم؟ ما هى أفضل إستراتيجية فى كتابة النص ؟. تركز الدراسة على تقييم أدوات تحويل النص إلى صورة المتوفرة بشكل مجانى (باستثناء *Mid journey* المعروف بقوة أداءه) باتباع الطريقة التجريبية والتحليلية والمقارنة والتي أدت للتوصل إلى بعض النتائج وأهمها ، الإستراتيجية المثالية لكتابة نص ناجح بالإضافة إلى إكتشاف *Microsoft Bing* كونه الأفضل على الإطلاق بالتجربة ، الأسباب التقنية وراء ذلك وغيرها الكثير .

الكلمات المفتاحية

الذكاء الاصطناعي؛ النص إلى صوره؛ التصميم الداخلي.

1. INTRODUCTION

There is no doubt that AI is starting to dominate the world every day more in all fields. A recent annual study by McKinsey on the state of AI and the rapid expansion of generative AI (gen AI) tools proves the theory and further ensured that there is no company or organization today not using AI tools in one form or another and it is increasing explosively every day where all those tools suddenly appeared in less than one year proving a rapid unstoppable industry (McKinsey, 2023) . The creative field is no exception. Recently the world has started to realize that AI tools are here to stay and we have to start thinking about how to use it for our benefit (being as interior designers or as academics guiding future designers as well). Believing in that future that will not be devoid of artificial intelligence, some researches especially in the academic field start integrating those tools in their teaching process and permit students using them .By deep investigation most of those trials are in the Architectural field (Author 2024), On the other hand almost little to no enough investigation to the usage of those tools in the interior design field in specific despite interior designing being a very challenging and a hard-work profession when it comes to the time consuming tasks related to ideation process or rendering and post production phase to the final design idea which can surely be solved using the current available AI tools making it a magical solution to all challenges facing an interior designer during the interior design process , but are those tools accurate enough for use ? . Can it really save time for designers reaching accurate results as accurate as designs generated by human expertise in the field (being knowledgeable of human centered design, styles and different design approaches ...etc.?). This study will focus only on investigating the accuracy of results generated from the text to image free available tools accuracy due to the increasing number of those every day. Accordingly this paper will try to answer the following questions divided into two main phases.

●Phase 1: (evaluating the accuracy of Text to image apps available) answering the following questions:

-Can Text to image apps generate accurate interior space solutions as required?

-Which is the most accurate available apps based on generated out comes (excluding the popular Midjourney)?

-What is the best prompt writing strategy for the most accurate results?

-Do those app performance improve over time significantly?

●Phase 2: Deeper investigation unveiling the potential of the app chosen as the most accurate in results answering the following questions:

-Does it understand different interior design zones / terminologies / design approaches?

-Can this tool be used as an assistant to interior designer in the design thinking process?

-Does the app generate applicable design solutions?

1.1 PROBLEM

-a pool of text to image out there, not knowing if they are reliable sources to help in the interior designing process with the same level of knowledge of a human expert.

 -a blurry vision on ai "text to image" real potential in generating accurate results as per required due to an unclear rule on how to write the best text prompt for the best outcome.

1.2. HYPOTHESIS

AI tools do generate good designs that can be at least used as an assistant speeding up the designing process and they are in continuous improvements, yet they cannot replace designers.

1.3. Aim

- Extracting the perfect prompt writing guidelines for the best results.
- Evaluate the accuracy of AI "text to image" generated interior space designs from a specialist's point of view applying the extracted and tested perfect prompt writing guidelines.
- Reach the best "text to image" app that can be used as a reliable source for the ideation step in the design thinking process by a designer.
- Deeper investigation of the (best) chosen AI app performance with different styles, design and interior spaces approaches for an accurate evaluation. (Future opportunities).
- Exploring if the chosen tool can really assist designers in the design thinking process.

1.4. Importance

-Evaluating the current popular AI tools that are here to stay and increasing every day and test the potential of using in the design process and level of reliability and accuracy of generated results by using the perfect prompt writing strategy reached based on experiment.

1.5. Methodology and Time Frame
-The paper will follow the experimental, analytical and comparative method. Experimental and analytical method by experimenting with different text to image tools and different prompt writing styles, analyzing the generated scenes as well as analyzing and deeply investigate the potential of the best chosen app (in regard to different interior design related zones / styles and design approaches).
-The comparative method is applied by comparing the results generated by different apps to conclude the best performance as well as comparing different prompt writing strategies impact on the generated design to conclude the best strategy on a span of almost a year (10 months).

1.6. Paper Structure

The paper is divided into 9 sections excluding the introduction.

Table (1) – paper structure (Author 2024)

Section title	Section content
2.AI tools in the creative field (background)	This section includes a literature review about currently available Ai tools in the field (classification and details).
3.Evaluating text to image AI tools in interior design	In this section evaluating different AI tools using a fixed prompt took place to conclude the best one (In 2023) .
4.Testing best prompt writing strategy	To further investigate selected Ai tool capability, testing with different prompts took place in order to conclude the best writing strategy.
5.Evaluating AI tools performance in a span of almost a year (10 months)	In this section the previously Top scoring AI tools are tested once again after months to make sure of tool consistency and accurately evaluate performance by time.
6.Evaluating best consistent AI tool deep understanding of Interior design	In this section, The selected consistent AI tool is tested regarding deep understanding of interior design (principles and approaches and terminologies).
7.AI as Human expert assistant	Exploring the chosen best AI text to image tool performance in designing process of a booth design (as an example).
8.. Results discussion and investigating deeper in reasons behind the results.	

9. Final Conclusion, recommendations and references.

2. AI TOOLS IN THE CREATIVE FIELD (BACKGROUND)

Driven by the rapid advancement of intelligent technologies, artificial intelligence has transitioned from the domain of science fiction to an indispensable component of modern life and when it comes to AI in the creative field (interior design included) , there are a myriad of tools that have surfed the internet that help in generating designs in a matter of seconds , open for use not only from professionals but from non-specialized people too, making it even hard to choose the best option for assistance even for professionals . All this came with the rapid development and popularity of what we call "Generative AI" that has proven to be a great source of creativity support in recent years (Michael et al., 2023). In this aspect, Generative AI tools available can be generally classified as follow (figure 1) :

Figure 1, classification of AI tools available in the creative field (Source: Author, 2024)

First: Text to Image

This is based on transforming input text description into realistic images. This type is based on a previously trained machine with a huge data set (text-image paired data), the more accurate the data sets are, the more accurate are the outcomes. This type works using convolutional neural networks (CNN) to convert extracted information into images as well as Natural Language Processing (NLP) to help the machine understand the text input given to it and transform it into images (Vaswani et al., 2017).

Second: Image to Image

These tools leverage deep learning techniques to translate one image into another, opening possibilities for editing, style transfer, object manipulation, and much more.

Third: Sketches to Realistic Images

This is considered the most recent advancement in deep learning, allowing users to transform simple sketches into detailed and realistic images relying on what is called Generative Adversarial Networks (GANs) (Goodfellow et al., 2014). Most popular model used in this type of AI is the diffusion model used to refine noise into a realistic image while learning from the sketch (Ho et al., 2020). This paper will focus on "Text to image AI tools and evaluate the performance on a span of almost a year (10 months).

3. EVALUATING TEXT TO IMAGE AI TOOLS IN INTERIOR DESIGN

3.1 Criteria of Selection

The study focused on evaluating popular text to image AI tools that are totally free and available to all audience (excluding the popular Midjourney that was partially free in early 2023) and for the

evaluation to be accurate the performance of the selected tools in the experiment are tested on a span of 10 months from the year (2023-2024) to decide which of those apps are getting better with time in performance. Particularly in June 2023, Text to image tools selected for testing are as shown in (figure 2)
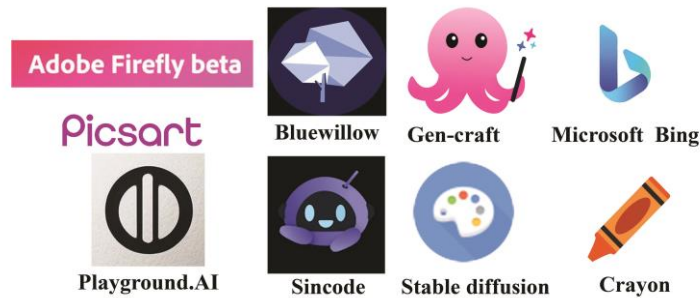

Figure 2, Text to image AI tools tested in 2023 (Source: Author, 2024)

3.2. Testing and Evaluation Process of AI Tools

To fairly evaluate the selected apps, The prompt text used is fixed and written as follow:
"An office with 40% oak wood mainly in flooring and windows with, lower ceiling height above the desk area which should be of oak wood with the top mainly of Carrara marble. Blue, beige color as a general color - scheme in an overall minimalism style".

Results from experiment (Author 2024) :
●First : Adobe firefly Beta : Not all the prompt description is applied accurately in the generated shots .There were serious technical errors in the generated images (e.g. : Some furniture pieces , circulation …etc.).
●Second: Blue willow : By analyzing the shots , it is clear that it has better , more real scenes with lesser mistakes than adobe fire beta with better treatments and textures , yet parts of the prompt where not applied accurately .
●Third: Gencraft: Over roll good out come with high resolution results, yet not accurate in some details, circulation and orientation of items in space are not right as well.
●Fourth: Microsoft Bing : The generated shots are to a feat extent aligned with the prompt (not fully) of great resolution and realistic out comes, yet there are some simple errors in some furniture pieces but no major technical errors or circulation errors. (The generated shots are totally functional) as opposite to some previous apps that generated serious errors in the matter.
●Fifth: Piscart: Only one shot of the generated was actually an office, while the other generated shots where mostly workstations. Over roll not the best outcome.
●Sixth: Playground.AI : The generated shot was not an office in the first place. (Wrong prompt translation)
●Seventh: Sin code: This AI tool had several modes (Dale 2 –open journey – stable diffusion mode) . The shot generated with stable diffusion mode was very good in applying materials as required, yet the circulation part was illogical.

●Eighth: Stable diffusion : Except for one shot, the generated shots were not a manager office in the first place, and the one generated had a myriad of technical errors.
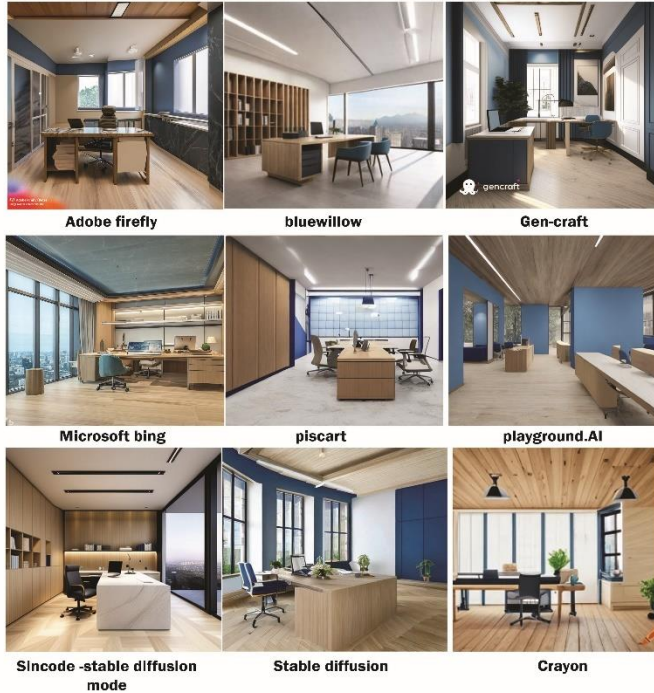

Figure 3, One generated shot from each app using the prompt (Source: Author, 2024)

To Conclude the best results is generated from which AI tool , a check list was done to check items achieved in the generated image from the prompt summarized as (percentage of wood in space ,

Figure 4, Analyzing apps outcome Correspondence to prompt text structure (Source: Author, 2024)

Oakwood in flooring and around openings , color palette and materials accuracy , marble and wood in surfaces touched by client and minimalism style as shown in (figure 4)

3.3. Findings

Based on the checklist, it was concluded that blue willow was absolutely the best option out of all tested AI tools in successfully achieving most of the prompt in the generated images for the year 2023, followed by Microsoft Bing and Sincode. To check Blue willow capabilities furthermore and if results can even get better, experimenting with prompt writing took place to reach the best prompt writing strategy and conclude fixed guidelines to use.

4.TESTING BEST PROMPT WRITING STRATEGY

The testing process took place through (figure 5):

Testing prompt strategy

| Paraphrasing by expert | Paraphrasing using AI tools | Paraphrasing using both | Paraphrasing using Google bard | Paraphrasing prompt generators |
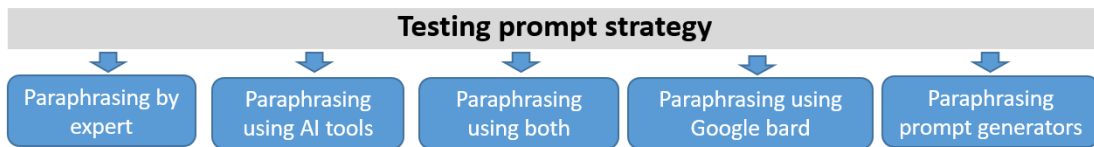
Figure 5, ways practiced in testing prompts impact on generated images (Source: Author, 2024)

As per figure (6) testing with prompts writing using all the five methods took place as follow:

●First: prompt by expert

"A minimal style office with 40% oak wood applied in floorings and around windows. There is a lower ceiling area above the desk area. The desk design be mainly of oak wood with a Carrara marble top. The general color palette should be of blue and beige".

●Second: paraphrasing using AI tools

"The office features oak wood, comprising 40% of the flooring and windows. The desk area has a lower ceiling height and is constructed mainly with oak, wood, while the top surface is predominantly made of Carrara marble. The overall color scheme of the office is characterized by shades of blue and beige, reflecting a minimalist style."

 ●Third: paraphrasing using both (expert and AI tools)

"A workspace that has 40% of its composition as oak wood, primarily utilized for the floor and windows, along with a reduced ceiling height over the desk zone. The material should predominantly be oak wood, and the top should mostly be composed of Carrara marble. The

Prompt by expert   Prompt using AI   Prompt using both
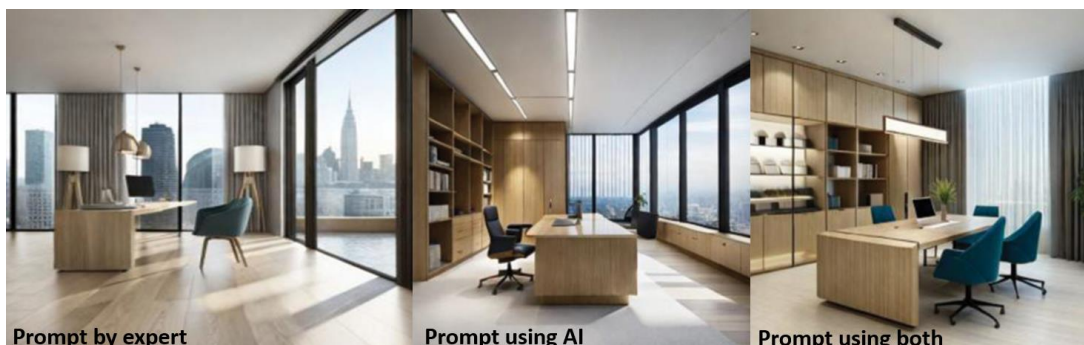
overall color scheme should be shades of blue and beige. Plan in a generally minimalist design style.”

Figure 6, shots generated using different prompt writing strategies (Source: Author, 2024)

Discussing Results: The generated images from the three prompts are not too far , yet the prompt generated using both expert and AI generated slightly better results as shown in (figure 5,p.7).

Further investigation took place trying to make a more detailed descriptive prompt of what was required and test the results. The prompt was written as follow and the results are shown in (Figure 7): “A three-wall manager's office scenario with a window in each of the two side walls (left and center), and the third wall (on the right side) with simply paint and a three-seat coach in front of it, and a white door at the end of the same wall. The window is located in 1/3 of the center wall, with a white door on the opposite side of the wall (from the left side of the wall) that connects to the office bathroom. In front of the same wall is the office desk, with two front chairs for customers. The office desk is constructed from oak and Carrara marble. Oak wood should be used to cover an 40% of the office space including the flooring and around windows area , while the

Figure 7, Generated shots from detailed descriptive prompt (Source: Author, 2024)

rest of space has the color palette blue , beige and white in an overall minimal style”.

Discussing Results : The main target of writing a descriptive prompt is to test if AI tool can generate rendered scenes to designer’s thought and in mind design (and not as a source of inspiration) . The results prove that the longer and more in details prompt , the more are the errors. By checking the generated shots , it is clear that some generated shots are not a manager office in the first place .

As per using Google bard, a not so short prompt was used as well and tested. The prompt is not a wall per wall description but a little bit longer prompt than the original prompt used at the very beginning. An example for prompt writing for one of the trials was as follow:

“An office with 40% oak wood, mainly in flooring and windows, with a lower ceiling height above the desk area which should be of oak wood with the top mainly of Carrara marble. Blue and beige colors are used as a general color scheme in an overall minimalist style. The flooring is made of light oak hardwood. The windows are also made of oak wood, with white mullions and frames. The ceiling height is lower above the desk area, creating a more intimate and focused space. The desk is made of solid oak wood, with a Carrara marble top. The chair is also made of oak wood, with a blue fabric seat. Other furniture in the office includes a credenza, bookshelves, and a sofa. These pieces are also made of oak wood, with beige upholstery. The overall color scheme is light and airy, with pops of blue and beige. The blue is used in the rug, pillows, and artwork. The beige is used in the upholstery, walls, and ceiling” .

Figure 8, Generated shots from Google bard (Source: Author, 2024)

Discussing Results : By analyzing results generated , Overall very good results and it is evident that The first 2 sentences in the prompt are always clearly reflected , after that not necessarily (it's a hit or miss). Each image included some of the requirements but not all .

For further investigations, Using different paraphrasing tools were tested as well. Examples to paraphrasing tools used are shown in Figure (9) .

Figure 9, Some tested paraphrasing tools to prompt (Source: Author, 2024)

Results were not the best as well. Last test was to order Google bard to write a perfect prompt for blue willow (as written in Midjourney) and giving the required description . The generated prompt resulted in the best results out of all previous trials (Author 2024) as shown in figure (10), which made the author analysis the prompt structure deeply as shown in figure (11)

Figure 10, Images generated ordering Google bard for a good prompt (Source: Author, 2024)

A minimalist manager office with a calming color palette of blue, beige, and white, and generous use of oak wood. The walls are painted a deep blue, the flooring is made of oak, and the window trims are also oak. The desk area is located in the center of the room, and it is made of oak wood with a Carrara marble top. The desk is surrounded by a few comfortable chairs, and there are a few other pieces of furniture in the office, such as a sofa, a coffee table, and a bookcase. These pieces are also made of oak wood, and they are kept simple and understated. The lighting in the office is natural and indirect, with large windows and light-colored curtains. The artwork in the office is minimal and abstract, and the plants are low-maintenance and easy to care for.

Figure 11, Analyzing the winning prompt structure   (Source: Author, 2024)

Conclusion : The best prompt structure should start with the style and type of space , followed by the general color palette . Mention the interior element / item and add a brief description each at a time (2 or setences mostly) . The shorter the better. As suumarized in the following figure (figure 12)

Figure 12, Concluded prompt writing structure (Source: Author, 2024)

5.EVALUATING AI TOOLS PERFORMANCE IN A SPAN OF MONTHS

To confirm Blue willow being the winner AI tool, the experiment was redone after 10 months to measure the performance (decreasing / getting better or as it is) . Using the original very first prompt before even reaching the perfect prompt writing strategy to compare the results. It was found that Blue willow became acquired by Lime wire and Blue willow became an internal model. The interface was slightly changed. And by testing the blue willow model and comparing it with the previous trials. It was found that the performance decreased and style errors where even generated as opposed to before , same goes with Sin code stable diffusion .Piscart was of same level while Microsoft Bing performance is getting better (Author 2024).

Figure 13, Images generated by Limewire bluewillow , Microsoft bing and sincode 2024 (Authur2024)

In 2024 , popularity to other AI tools have increased (e.g: Lexica and Leonardo.AI) ,when tested using the same old prompt (before the right structured one to compare the results) , the outcome was overall good , yet the prompt was not accurately achieved and by repeating using a well- structured prompt (based on previously extracted guidelines by author), the overall results to all three was better, yet Microsoft Bing generated the best results .

Figure 14, Images generated following prompt writing extracted guidelines (Authur2024)

Results and conclusions (Author 2024):

-Using the concluded prompt writing guidelines lead to overall better results and Microsoft Bing proved to have consistent results and getting better on all the experiment span as compared to other AI tools from 2023 as well as 2024.

-The overall results were very satisfying leading to experimenting more with the AI tools to conclude extent of Microsoft understanding of different interior design (principles / approaches / styles and even different interior spaces).

6. EVALUATING MICROSOFT BING DEEP UNDERSTANDING OF INTERIOR DESIGN

To deeply test Microsoft Bing understanding of Interior design. The testing process will cover different aspects as summarized in (figure14)

Figure 15, experiment structure for Microsoft Bing understanding (Authur2024)

6.1. Testing Microsoft Bing Understanding of Interior Design Principles

In this section, each point under each category was tested separately following the prompt guidelines previously concluded. Regarding the terminologies section, 3 of interior design principles were tested (Asymmetrical , radial balance and contrast) . The experiment showed that that some terminologies are understood and rightly applied (e.g: Asymmetrical) , but other errors were generated regarding functionality and circulation despite generating overall very good images with a lot of details . Other terminology like (radial balance) , Microsoft Bing failed to achieve it , leading to either translate it into a circular wall treatments (which is wrong) or trying to generate circular sitting configuration was a myriad of technical errors , when it came to contrast , No contrast in color palette was achieved despite generating very good images (Figure 16) .

Figure 16, Images generated for testing design principles following guidelines (Authur2024)

6.2. Testing Microsoft Bing Understanding of Styles

As examples (Victorian , art deco and industrial styles were tested) , the results are overall very satisfying and the styles were understood , yet in Victorian style 2 of the three images were clearly illustrations and not realistic images .
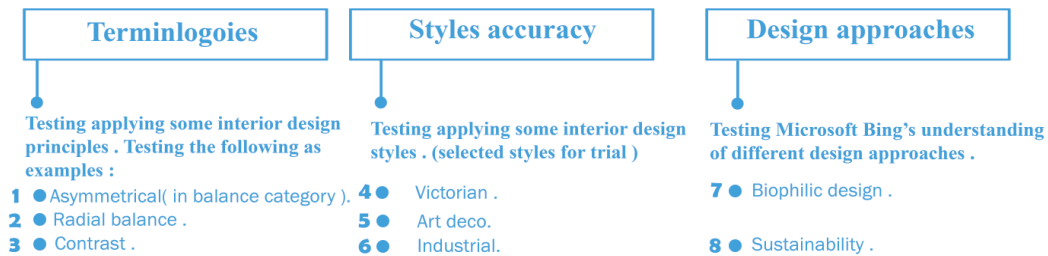

Figure 17, experiment structure for Microsoft Bing understanding of styles (Authur2024)

6.3. Testing Microsoft Bing Understanding of Different Design Approaches

In this section, different design approaches (e.g :Biophilic and sustainable design were tested) firstly on an office space then those approaches were tested as well on different type of interior design spaces which the author classified as follow (figure 18)

Figure 18, classification of different interior spaces   (Authur2024)

Results: Microsoft understood all spaces overall , when it came to different design approaches , sustainability and biophilic design were translated in generated scenes as same (usage of good amount of wood and some greenery) ,which is not completely wrong.

7.AI (MICROSOFT BING) AS HUMAN EXPERT ASSISTANT

In this section, A trial will be conducted utilizing Microsoft Bing in the designing process.
-Target: Designing a booth design for a company specialized in packaging , printing and branding called creative monkey founded in the year 2006. This experiment will be divided into two phases as follow:
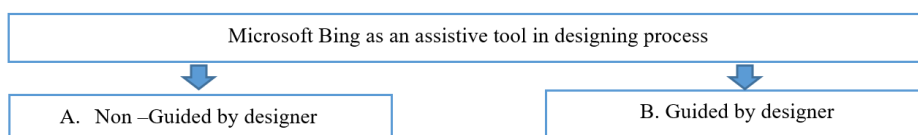

Figure 19, classification of Microsoft Bing usage as an assistive tool (Authur2024)

A.Non- Guided by designer: Utilizing Microsoft Bing as an ideas generator without being directed by designer and evaluate the results . In this regard, it is worth mentioning the booth types in general can be classified as (In line booth – corner booth –peninsula and Island booth) . The first stage of the experiment is to use Microsoft Bing as a designer assistant to generate a design for each type following the previous extracted prompt writing guidelines.

Trial example: "A(line booth type design) for a packaging industry related company .The booth design should be made of sustainable materials. It shall include a small meeting table , space for circulation and displaying areas" .Same prompt was used several times by replacing inline type with other types .

Observation: Except for the corner booth design , Microsoft Bing did not understand other booth types terminologies , yet the outcome generated shots were good and applicable to a great extent and other parts of the prompt are achieved in the generated scenes as well. Another trial was done replacing booth type terminology (not understood by Microsoft Bing with a description of it (e.g.: using the same prompt but replacing inline booth with a booth closed from 3 sides and open from one side) , yet no accurate results were generated .

Figure 20, Microsoft Bing usage as an assistive tool in booth design (Authur2024)

B. Guided by designer:

In this trial, An idea was in the researcher mind based on some pre-studies related to the company's background and an idea by the researcher to what color palette and materials to use , which was chosen (as black and whites mainly) and corrugated carton as an optimum sustainable material making a statement and reflecting the company's activity in a clear way . Firstly, The researcher tried to check to check if Microsoft Bing can generate a mood board by directly ordering this through a directed prompt by researcher as follow : " A mood board for a company specialized in packaging , branding and printing. Corrugated carton should be the main material used". The results were not mood boards at all, but shots near to previously generated shots (Figure 21 p.14). Another trial was conducted using a directed prompt as follow : "A booth interior with corrugated carton material mainly used .The booth should be a formation of original colored corrugated carton boxes and black ones with text graphics for a company specialized in packaging , printing and branding" .(figure 22 p.14). By trying to be even more specific another directed prompt mentioning the company's name was stated as follow : "A booth interior with corrugated carton material mainly used . The booth should be a formation of original colored corrugated carton boxes and black ones with text graphics for a company specialized

in packaging , printing and branding with the name creative monkey (of black and white logo)”. (figure 23).

Figure 21, Microsoft Bing generating a booth design not a mood board (Authur2024)

Figure 22, directed prompt by researcher results (Authur2024)

Figure 23, directed prompt by researcher results including company's name (Authur2024)

Conclusion by researcher:

Microsoft bing did not understand some core interior terminologies (e.g : Mood bard) , yet the directed prompt strategy applied in (Figure 22 and 23 resulted in overall good results that can be used as they are by only customizing dimensions and were used as an inspiration by designer .

7.1. Conceptual Design Proposal by Researcher

The main concept was specified by the researcher as follow : "Crafting Identities with Shared Sustainability" , meaning saving the company's identity and reflecting their different activities by using sustainable materials that were already used in booth before but in a new creative way that match the company's identity using corrugated carton as the main material .

Figure 24, proposed conceptual elevation design for in booth design (Authur2024)

Figure 25, proposed conceptual shot design for in booth design (Authur2024)

The conceptual design was generated for two booth types (Inline booth being the most popular and the most used and the island one being a little distinct than the other two types) where figure 22 and 23 p.14) were a good source of inspiration .

The main idea was to utilize corrugated cartoon box and material in a creative way and playing with black and white palette to match the company's logo and general theme and adding a catchy pop of color through pink to grab attention with decorative stripes on boxes and graphic prints (Generated using Microsoft Bing). In the inline booth design , it started with designing the main elevation as this is the main view for this type of booths shown in figure 24 (p.14) . Another quick option for the island type was proposed (figure 26 and 27)
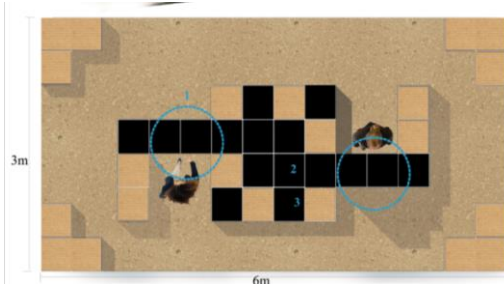

Figure 26, proposed conceptual plan design for island booth (Authur2024)
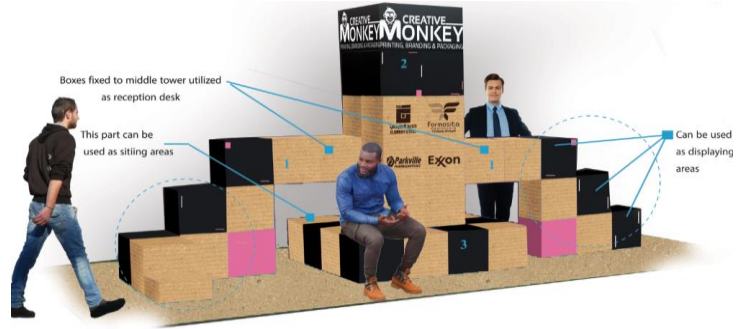
Figure 27, proposed shot design for island booth (Authur2024)

8.RESULTS DISCUSSION AND DEEPER INVESTIGATIONS ON REASONS BEHIND IT

8.1. Results Discussion:

8.1.1. Regarding AI (Text to Image Tools) Evaluation

-Most of the text to image tools generated different errors (either technical / circulation or non-understanding for some interior space terminologies).
-BlueWillow was one of the best Text to image tool in 2023 , while Microsoft Bing proved to be more consistent by not only maintain same level of generated interior scenes but improving by time till now making it the best option (and even overtaking BlueWillow after merging with lime wire in 2024) .

8.1.2. Regarding the Optimum Prompt Writing Strategy

The best prompt writing strategy is to start with the type of interior space, style (if needed) , color palette , then start with the item name and description related . (The shorter the better) and being more directed from designer's point of view.

8.1.3. Regarding Using AI as an Assistive Tool in the Designing Process.

Microsoft Bing do not understand a lot of interior space terminologies. Yet, it is a reliable source to be used as an assistant in the designing process .It can help in generating inspirational ideas that are very good and applicable as well yet the unique touch of each designer will be lost. While by utilizing it as an assistant to performs designer's imagination and point of view by specifying certain requirements (specific materials, color palettes ...etc.) , it generated outstanding designs that can be applied directly. It just needs to be customized as per the actual dimensions of the actual interior space and using a directed prompt strategy for the best results (designer being the designer and Microsoft Bing being the tool).

8.2. Deeper Investigations on Reasons Behind those Results

The results concluded pushed the author to investigate further and deeper in the reasons that could have contributed into the errors generated by some of the text to image tools investigated and why some are better than others and why Microsoft Bing came out the best. The Author fears we're becoming slaves to these tools if we became only users, the fact that lead deeper investigation on the technical issues behind the matter a must to build accurate final conclusion and recommendations. Luckily, some (non-designer) background researchers mostly from all over the world believed in this as well which reflected into the numbers of papers produced recently and increasing every day for the past few years on popular scientific platforms like IEEE Xplore , Google scholar and arXiv focusing on improving current available Text to image's performance by improving related models with very little focusing on Text to image apps related to interior design unfortunately (Authur 2024).Yet , That does not deny that the papers author found can give us a good insight and overview evaluation of the current AI text to image tools and performance which surely have improved a lot since its introduction in 2014 and is continues improvement where generative diffusion models, including DALL E, Midjourney, Stable diffusion, and others and further more are introduced by developers every day for even better results (Alhabeeb & Al-Shargabi, 2023) and by investigating the little comprehensive papers targeting improving text to image apps for better interior shots generated in specific , valuable improvement suggestions were found .

One of the most prominent papers in the matter was a paper produced in the year 2024 that focused more on interior design domain by producing a model called "I designer" where the owners teamed up with a world-famous interior design company. This partnership gave them access to a massive historical archive of designs, created by over 1,000 design professionals. This archive was used to build a high-quality dataset of 3,600 image-text pairs, where each image has a detailed description (Smith, J.2024). The model was based on stable diffusion XL (SD-XL) which is the most popular T2I open sourced model in the world based on experts (Dustin, Zion, et2023) and as a part of the paper the authors compared their proposed model to other popular ones with the model achieving a win rate of over 58% against both DALL-E 3 and SD-XL base lines. The model as well took into consideration language barrier in writing prompts leading to designing it as bilingual text to image model (English – Chinese). Another recent paper (Chen, Shao, et 2023) suggested a new diffusion model and throughout the process, it compared their model to other popular ones (Disco diffusion , Midjourney , DaLL E2 and stable diffusion) , regarding items very aligned with interior space mentioned in the table below (Table 2 , p.16) . By excluding Disco diffusion and evaluating the other popular and more specified models excluding their model as well for now.

Table 2 _ summarizing popular model performance in interior design items (Author 2024)

Point of comparison	Midjourney	DaLL E2	Stable diffusion
Style	Best	lesser	Least
Functionality	Best	Of same level (lesser than mid journey)	
Furniture positioning	Lesser	Least	Best
Object integrity (Accurate space to function)	Best	Lesser	Least
Design details	Best	Least	Lesser
Realism	Lesser	Best	Least

Based on the table above, it is evidence that Midjourney is the best, that's why it was excluded from the beginning of the paper based on its known performance as well as being non free.

While between DaLL E2 and stable diffusion, DaLL E2 is slightly better than stable diffusion overall with more positive items related to interior being better than the latter in style, object integrity and realism (Author 2024). when going back to the apps evaluated (figand trying to investigate which model it belongs, most of the apps' diffusion models are proprietary . The only known are Sincode stable diffusion and playground.ai which of course uses stable diffusion (open source) , Leonardo.ai and Lexica are known to use stable diffusion as well as other un known models while Microsoft Bing (Co-pilot) that was proved to be the best one by experiment uses DaLL-E and a mix of other Open AI models (Author 2024) .

9. FINAL CONCLUSION AND RECOMMENDATIONS.

9.1. Conclusion

-Throughout the author's own experiment, it is evident that Text to image tools have a great potential to be a great assistant to designers when utilized rightly, this was further proved by recent researchers finding ways to improve performance by updating diffusion models.

-Following the extracted prompt writing guidelines by Author is a must to get the best results.

- Text to image apps based on DaLL-E model give better performance when it comes to interior design and mixing it with other models improves the performance and that justifies Microsoft Bing being the best out there and other apps based on DaLL-E model (example: Lexica, and Leonardo.AI) generating acceptable solutions .

-Currently, diffusion models struggle to capture all the details from a text description, even though they are trained on data combining text and images which is the main reason for the errors produced even by the best app there (Microsoft Bing / Copilot) , the fact that was clearly addressed in a recent paper evaluating models as well (Cao, Zhou 2024).

9.2. Recommendations and Future Studies (Author 2024)

-Interior designers and computer scientists should work together hand in hand as interdisciplinary work to improve text-to-image generation models used in interior design by contributing in accurately labeling data based on scientific knowledge about right terminologies, style characteristics, and functionalities from a professional point of view while using DaLL-E model as one on the ones used.

-working on improving Microsoft Bing (Copilot) performance in specific being the best out there and in continuous improvements by working on labeling data more accurately (as previously mentioned) as well as working on other options or services that can enhance using the app as an assistive tool without taking control on designer . For instance, it can have an option of modifying on generated output interior images through additional prompts and real time customization and interaction. It can have an option of adding a semi rendered image / extruded shot by another app of designer's actual work and using Microsoft Bing as a modifying tool using prompts to do the required or use a ready image as an inspiration which is a close idea to what was already exposed in a recent published paper which is a novel pipeline that generates personalized outputs based on exemplar images, not text prompts and has proved to generate amazing results (Xu, Guo,et2023) .That in addition to adding dimensions options and train the model to read very detailed prompt including dimensions , opening places ..etc.

-All in all and Generally there are some areas that need more investigation on how to implement in text to image apps in the future. The most prominent to be: Integrating functional aspects like

ergonomics, accessibility and universal design, Real-time customization and interaction , last but not least , Cross-modal retrieval combining text, images, and 3D models (that can be even generated by designers on other platforms) to generate another designer's idea at the end and ensuring designer taking control .

-From an academic perspective and by the appearance of a myriad of AI tools that are accessible to all including students anyway, Further investigations on how to use Microsoft Bing (Copilot) in the teaching process of Design studios should take place, but before it an awareness for instructors about AI tools opportunities and capabilities should take place to integrate the right way .

10. REFRENCES:

Alhabeeb, S. K., & Al-Shargabi, A. A. (2023). Text-to-Image Synthesis with Generative Models: Methods, Datasets, Performance Metrics, Challenges, and Future Direction. *IEEE Access*. Advance online publication. DOI: 10.1109/ACCESS.2024.3365043

Cao, P., Zhou, F., Song, Q., & Yang, L. (2024). Controllable generation with text-to-image diffusion models: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence. Retrieved from https://ar5iv.labs.arxiv.org/html/2403.04279

Chen, J., Shao, Z., & Hu, B. (2023). Generating Interior Design from Text: A New Diffusion Model-Based Method for Efficient Creative Design. *Buildings*, 13, 1861. https://doi.org/10.3390/buildings1307186

Dustin Podell, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas M¨uller, Joe Penna, and Robin Rombach. Sdxl: Improving latent diffusion mod els for high-resolution image synthesis. arXiv preprint arXiv:2307.01952, 2023

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks [arXiv preprint arXiv:1406.2661]. Retrieved from http://papers.neurips.cc/paper/5423-generative-adversarial-nets.pdf

Ho, J., Jain, A., Abbeel, P., & Malik, J. (2020). Denoising diffusion probabilistic models. In H. Larochelle, M. Balcan, R. Fergus, S. Bachman, D. Tran, J. Hinton, & E. Xing (Eds.), *Advances in Neural Information Processing Systems 33* (pp. 18242-18256). Curran Associates, Inc. https://arxiv.org/abs/2006.11239

McKinsey & Company. (2023). The state of AI in 2023: Generative AI's breakout year. McKinsey & Company- https://www.mckinsey.com/~/media/mckinsey/business%20functions/quantumblack/our%20insights/the%20state%20of%20ai%20in%202023%20generative%20ais%20breakout%20year/the-state-of-ai-in-2023-generative-ais-breakout-year_vf.pdf

Muller, M. (2023). Generative AI and HCI at CHI 2023 [Workshop]. In CHI '23: Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (pp. 1-6). Association for Computing Machinery: https://www.acm.org

Smith, J. (2024). iDesigner: A high-resolution and complex-prompt following text-to-image diffusion model for interior design. *Journal of Interior Design Research, 12*(3), 45-67. https://arxiv.org/abs/2312.0432

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In H. Larochelle, M. Balcan, R. Fergus, S. Bachman, D. Tran, J. Hinton, & E. Xing (Eds.), *Advances in Neural Information Processing Systems 30* (pp. 3015-3025). Curran Associates, Inc.

Xu, X., Guo, J., Wang, Z., Huang, G., Essa, I., & Shi, H. (2023). Prompt-Free Diffusion: Taking "Text" out of Text-to-Image Diffusion Models (Version 2). arXiv preprint arXiv:2305.16223v2.

Zhao, L., Yuan, Z., Zhang, Y., Wu, S., & Wu, J. (2024). Learning Generative 3D Scene Layouts from a Single Image- CVPR 2024 Workshop on AI for 3D Generation (conference paper) - (1) (PDF) Learning Generative 3D Scene Layouts from a Single Image (researchgate.net)