Enhancing Design Ideation with AI: Integrating LLMs and VLMs into Industrial Design

AIM: To investigate and legitimize the integration of Large Language Models (LLMs)* and Vision-Language Models (VLMs)* into the early design ideation processes, exploring effects on creativity and efficiency in industrial design.

METHODOLOGY

- Developed a between-subject design, recruited participants, and conducted the study with a sample size of 35 participants (N=35)
- Determined contextually feasible design cases within brainwriting ideation method* and video-based design (VBD)*
- Collected quantiative and qualitative data: workload (NASA-TLX), experience (UEQ), acceptance of technology (UTAUT), interviews and eye-tracking technology

KEY FINDINGS

AI Can Support Ideation:

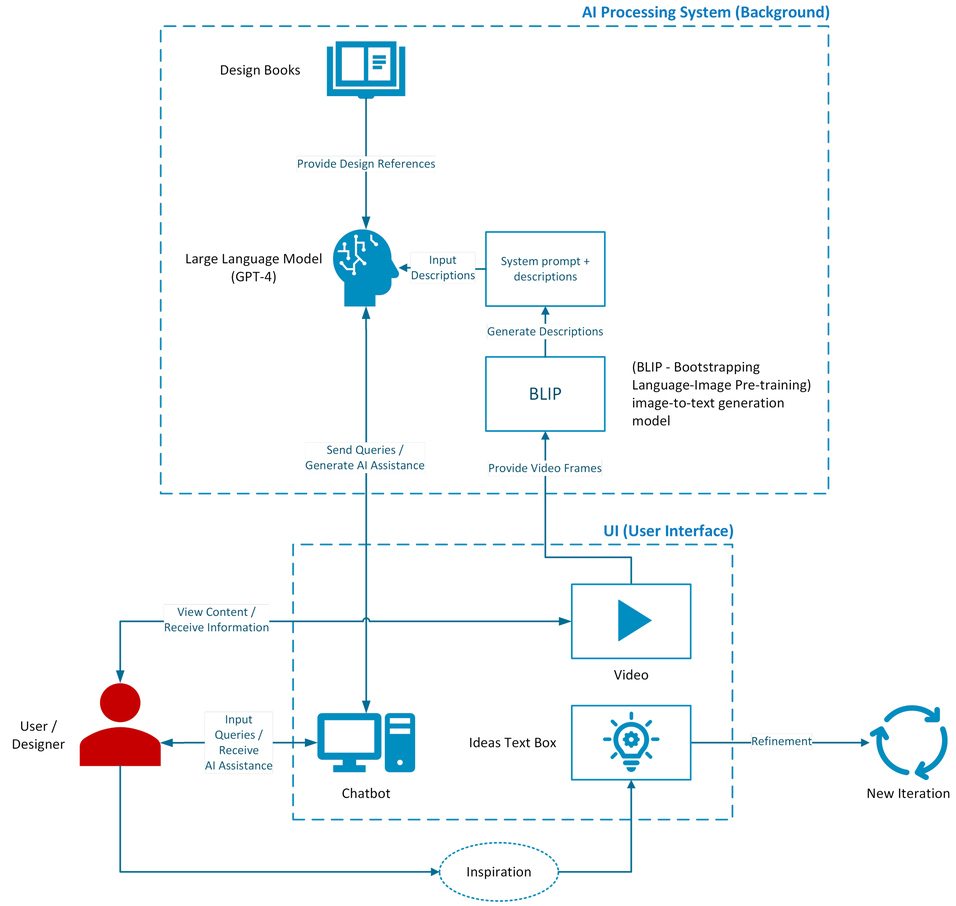
- AI has potential to generate a broad range of ideas and designers view it positively to handle repetitive tasks.
- Requires improvements for contextually relevant and accurate information.

Cautious optimism:

- Designers are cautiously optimistic about AI's potential.
- Emphasis on the need for human oversight to retain creativity.

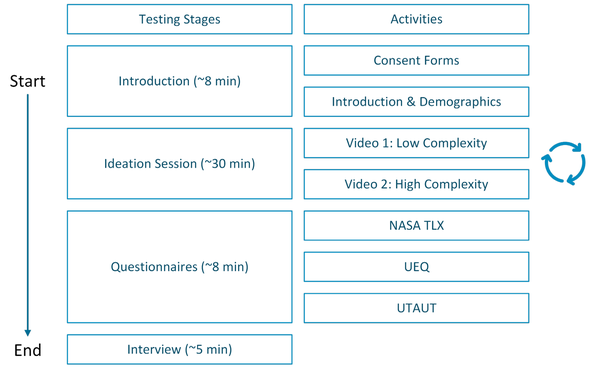
AI as a Collaborative Partner:

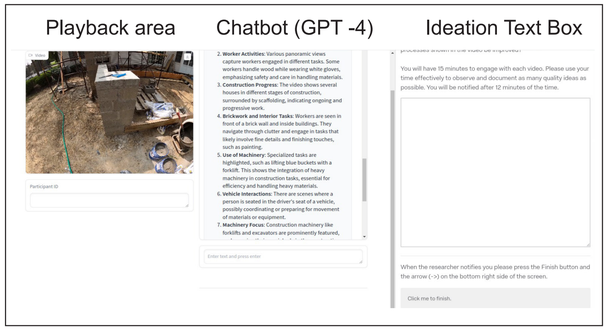
- AI should function as a collaborative partner, not a decision-maker.
- Necessary to preserve the designer's role as the critical actor in the creative process.

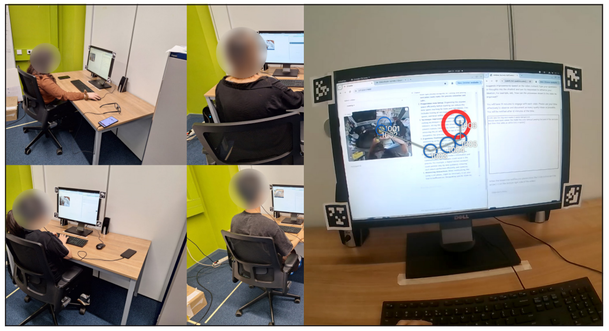


Interaction Flow

- **BLIP**: Processes video frames for contextual understanding.
- **LLM (GPT-4)**: Receives descriptions and generates potential ideas.
- **UI (Chatbot)**: Input queries and receiving AI-generated ideas.
- **Ideas Text Box**: The output of the brainwriting ideation process, where AI aids in generating ideas.
- **Design Books**: Sources of information integrated into the AI for generating contextually relevant ideas.







- *Brainwriting is a specific ideation approach for generating a great number of ideas. Usually performed with individuals writing down their ideas on their own. Developed to speed up the idea generation and based on the assumption that quality leads to quantity.*
- *Large Language Models (LLMs): A probabilistic model that assigns a probability P[w1,w2,...,wn] to every finite sequence w1, ...,wn (grammatical or not); trained on wide amounts of data, making them capable of understanding and generating natural language and other types of content to perform a wide range of tasks.*
- *Vision language models (VLMs) are multimodal models which use both visual and textual data to perform tasks that require understanding of both modalities. These models are trained on datasets containing images paired with descriptive text, allowing them to generate descriptions for images, answer questions about visual content, and even generate images from textual descriptions.*
- *Video-based design (VBD) is an approach that integrates video into the design process, enabling a deeper, user-centered understanding by capturing real-life contexts and experiences to inform and enhance design outcomes ("Studying what people do," 2007)*

REFERENCES: Speech and language processing. (n.d.). https://web.stanford.edu/~jurafsky/slp3/ / Vision Language models explained. (n.d.). https://huggingface.co/blog/vlms / What is Brainwriting. (2024, March 22). The Interaction Design Foundation. https://www.interaction-design.org/literature/topics/brainwriting / Ylirisku, S. P., & Buur, J. (2017). Designing with Video: Focusing the user-centred design process. Springer.

Andrija Stanković
Understanding Design Ideation with Vision-Language Models and Video-Based Design
22.7.2024.
Integrated Product Design

Committee Dr. Evangelos Niforatos
MSc Tianhao He