UNIVERSITY OF TARTU
Institute of Computer Science
Conversion Master in IT

Mattias-Silvester Mustonen

AI-enabled prototyping tools in user-centered design

Master's Thesis (15 ECTS)

Supervisors: Grace Eden, DPhil

Yana Halas, MSc

Tartu 2025

AI-enabled prototyping tools in user-centered design

Abstract:

Recent advancements in Artificial Intelligence (AI) and the rapid development of generative models have created opportunities to integrate AI into digital prototyping applications, thereby enhancing functionality and capabilities. This thesis examines the current landscape of AI-enabled prototyping tools and centers its investigation around three research questions:

RQ 1: "*What is the current landscape of AI-enabled prototyping tools?*"

RQ2: "*How will users assess the quality and usability of an AI generated prototype?*"

RQ 3: "*How can an AI-generated prototype be iteratively improved based on participant's user feedback?*"

The findings indicate that the landscape of AI-supported prototyping tools remain in an early developmental phase, with only two applications—Visily and Uizard—demonstrating notable maturity in this domain. This suggests that these tools possess sufficient functionality and usability to be effectively employed in prototyping activities. An experiment was conducted using Visily to create a prototype of the Tartu Bike Sharing application. Participant evaluations revealed that, although the tool was capable of generating a functional and visually coherent prototype based upon prompts, it exhibited shortcomings in critical usability aspects (see section 5.3). Refinement efforts informed by user feedback highlighted limitations in the prototype's editing functionality, as well as Visily's insufficient ability to incorporate more complex elements, such as city maps.

The primary contribution of this thesis lies in providing practical insights into the capabilities and limitations of AI in the context of digital prototyping. By identifying the scope of AI's potential, the required skill sets for their use in the design process can be further assessed, enabling opportunities to enhance efficiency and design quality. Given the rapid evolution of AI technologies, it is recommended that future research be conducted across different case studies.

Keywords:

Prototyping, AI Tools, user experience

CERCS: P170 - Computer science, numerical analysis, systems, control

Tehisintellektil põhinevad prototüüpimise tööriistad kasutajakeskses disainis

Lühikokkuvõte:

Viimased edusammud tehisintellekti (AI) vallas ning generatiivsete mudelite kiire areng on loonud uusi võimalusi, et integreerida AI digitaalsetesse prototüüpimise rakendustesse, parandades sellega nende funktsionaalsust ja võimekust. Käesolev magistritöö uurib AI-toega prototüüpimise tööriistade praegust olukorda ning keskendub kolmele uurimisküsimusele:

UK1: *„Milline on AI-toega prototüüpimise tööriistade hetkeseis?"*
UK2: *„Kuidas hindavad kasutajad AI-ga loodud prototüübi kvaliteeti ja kasutatavust?"*
UK3: *„Kuidas saab AI abil loodud prototüüpi iteratiivselt täiustada, tuginedes kasutajatelt saadud tagasisidele?"*

Töö tulemused näitavad, et AI-toega prototüüpimise tööriistade areng on alles varajases faasis. Ainult kaks rakendust—Visily ja Uizard—on näidanud märkimisväärset küpsust selles valdkonnas. See viitab sellele, et neil tööriistadel on piisav funktsionaalsus ja kasutatavus, et neid tõhusalt prototüüpimise tegevustes kasutada. Uuringu käigus loodi Visily abil Tartu rattaringluse rakenduse prototüüp. Osalejate hinnangud näitasid, et kuigi tööriist suutis luua funktsionaalse ja visuaalselt kooskõlalise prototüübi etteantud sisendite põhjal, esines olulisi puudujääke kasutatavuse võtmeaspektides (vt lõik 5.3). Kasutajate tagasisidest lähtuvad täiustused tõid esile piirangud prototüübi redigeerimisfunktsionaalsuses ning Visily vähese suutlikkuse kaasata keerukamaid elemente, nagu linnakaardid.

Töö peamine panus seisneb praktiliste teadmiste pakkumises AI võimalustest ja piirangutest digitaalse prototüüpimise kontekstis. AI potentsiaali ulatuse määratlemine võimaldab hinnata vajalikke oskusi nende kasutamiseks disainiprotsessis, luues eeldused efektiivsuse ja disainikvaliteedi parandamiseks. Arvestades AI-tehnoloogiate kiiret arengut, on soovitatav tulevikus läbi viia täiendavaid uuringuid erinevate juhtumianalüüside põhjal.

Võtmesõnad:

Prototüüpimine, AI tööriistad, kasutajakogemus

CERCS: P170 Arvutiteadus, arvutusmeetodid, süsteemid, juhtimine

Contents

1. Introduction

Digital prototyping applications have served as effective collaboration tools between designers and users for several years. With the emergence of AI-enabled prototyping tools, new opportunities for enhanced collaboration between designers and users have been realized, elevating the level of interaction and engagement. Traditionally, prototyping—particularly high-fidelity prototyping—has required specialized design skills and in-depth knowledge of the tools to produce prototypes suitable for user testing [1]. The introduction of AI-enabled digital prototyping tools has the potential to significantly improve the efficiency and accessibility of creating digital prototypes. However, it is crucial to establish appropriate expectations regarding the outcomes of AI-enabled tools, and this thesis aims to investigate precisely that. This thesis seeks to investigate the current state of AI-enabled prototyping tools, structuring its inquiry around three primary research questions (RQ). RQ1 to be answered is: "*What is the current landscape of AI-enabled prototyping tools?"* Based on RQ1 application—Visly[1]—will be selected for further analysis. This analysis will serve as the foundation for addressing the subsequent research question. RQ 2 is defined as: "*How do users assess the usability of an AI generated prototype?".* RQ 3 is defined as: *"How can an AI-generated prototype be iteratively improved based on user feedback?".* To address this, a participant evaluation comprising three tasks will be conducted and based on user's feedback improvements will be incorporated into the prototype.

Following the introductory chapter, Chapter 2 provides essential background information, offering an overview of key concepts related to generative language models, user-centered design, and prototyping. This chapter establishes the relevance and significance of artificial intelligence in the domain of prototyping.

Chapter 3 offers a synthesis of key findings from relevant academic literature, outlining the main contributions within this field of study.

Chapter 4 details the development of the research methodology. Through the application of predefined criteria and a systematic selection process involving 20 applications, the platform Visily—noted for its AI-assisted design generation capabilities and prompt-based interaction—was identified as the most appropriate tool for the practical experiment. This chapter describes the development process of the *Tartu Bike Sharing* app, a practical case study conducted using

[1] https://www.visily.ai/

Visily. Additionally, it outlines the predefined success criteria based on a three-tier metric, as well as the methodology for user evaluation involving five participants.

Chapter 5 presents the findings in response to the thesis' three research questions. It begins with an overview of the current landscape of AI-supported prototyping tools. Subsequently, it reports the results of the user evaluation conducted on the AI-generated prototype. Finally, user-generated suggestions for improvement are analyzed and integrated into the prototype.

Chapter 6 discusses the results and outlines the challenges encountered throughout the study and Chapter 7 provides the conclusion together with recommendations for future research directions.

2. Background

The following section and subsections provide an overview of concepts that are relevant to the area of research in the scope of this thesis.

User-centered design (UCD), a concept introduced by Don Norman in 1996 defines that user input on purposes for design artifacts is central to their success [12]. Users are put first in every phase of development, from brainstorming to testing [2]. User-centered design process is a project-oriented approach to interactive systems development, and it involves users, tasks, and goals to find decisions and requirements concerning developments. The UCD process helps to ensure the high usability of the artifact and is entitled as process, method or framework as well. Several modifications about the UCD process have been published, depending on the focus of the researcher but the fundamentals of the UCD process appear in the 4 step process described in Table 1 [11].

Table 1. User Centered Design (UCD) process.

Process step	Activity
Study/Analysis/Requirement	Appoints the user requirements, analysis user, context and system specifications that will be carried out the whole activity in the UCD process.
Design/Prototype	Refers to UCD process is design requirements or instructions of the system and users which would be developed throughout the UCD process.
Build/Development	Indicates the development activities of the systems or artifacts where requirements are gathered from the prior stages.
Evaluation	Ensures effectiveness and efficiency of the artifacts that meet user requirements.

Within the User-Centered Design (UCD) process, artificial intelligence (AI) is primarily utilized for generating design solutions and facilitating the automated development of these solutions [16]. AI plays a significant role in various tasks, including the creation of user personas, the evaluation of design solutions, the automation of repetitive tasks, and the enhancement of prototype fidelity [16].

2.1. Prototyping

The term prototype in software development refers to an initial version of a software system that is used to demonstrate concepts, try out design options, and find out more about the problem and its possible solutions [3]. Prototyping is an essential part of the User-Centered Design process. Prototyping enables designers to develop representations of interactive systems at an early stage in the design process. Designers and users can better communicate with one another when a prototype is used to demonstrate the system's features and interface [2]. Prototypes can be categorized according to various dimensions, including strategic purpose, development stage, and fidelity level. A range of prototyping methods corresponding to these categories will be discussed below.

Four distinct prototyping strategies can be implemented:

1. Horizontal prototype strategy develops one entire layer of the design at the same time. A user interface is created without any underlying functionality, making it possible to test how the user will interact with the user interface without worrying about how the rest of the architecture works.
2. Vertical prototype strategy is to ensure that the designer can implement the full, working system, from the user interface layer down to the underlying system layer. Vertical prototypes are generally high-fidelity software prototypes because their goal is to validate a particular idea at the system level, usually at the early stage.
3. Task-based prototype strategy organizes a series of tasks, which allows both designers and users to test each task independently, systematically working through the entire system. The tasks include only the functions necessary to implement the specified set of tasks. They combine the breadth of horizontal prototypes, to cover the functions required by those tasks, with the depth of vertical prototypes, enabling detailed analysis of how the tasks can be supported on a system level.
4. Scenario-based prototype strategy is similar to task-oriented ones, except that it does not stress individual, independent tasks, but rather follow a more realistic scenario of how the system would be used in a real-world setting. Scenarios are stories that describe a sequence of events and how the user reacts [10].

Different prototyping stages can be defined based on a given use case that describes the intended interaction between the user and a system. For a first demonstration of a concept visual prototypes are typically used that already represent the intended screen or interface designs of the concept but otherwise only simulate the system behaviour. Visual prototypes do not require any coding. As actual system behaviour can only be tested when at least some system functions have already been implemented, there is a need to develop the prototype further. This stage is defined as a functional prototype. In the case of a working prototype, implementation is so advanced that the application concept can be tested in field tests in real-world application scenarios [3].

Prototyping doesn't have to be complex and expensive, and they should command only as much time, effort, and investment as are needed to generate useful feedback and evolve an idea. The more "finished" a prototype seems, the less likely its creators will be to pay attention to and profit from feedback and therefore learn about the strengths and weaknesses of the idea and to identify new directions that further prototypes might take [4]. Therefore, it is essential to find the right prototyping method.

Fidelity describes how easily prototypes can be distinguished from the final product and can be manipulated to emphasize aspects of the design. Prototypes more similar to the final product are "high-fidelity" while those less similar are "low-fidelity." A high-fidelity prototype is often made with the same methods as the final product and hence has the same interaction techniques and appearance as the final product. Low fidelity prototypes, such as sketches, differ from the final product in interaction style, visual appearance, and/or level of detail [18].

Different prototyping methods apply to different fidelity stages of prototyping. In the following list, starting from low-fidelity and moving higher, different methods are described.

- A sketch is a freehand drawing on paper that gives a really low-fidelity representation. It is a fast way to get an idea ready for brainstorming and can describe it better than words.
- A storyboard is often a series of roughly drawn comic book-style frames that visualize key moments in a design and can be used to build a short narrative. This method can help in fine-tuning an idea and provide insight in who will use it, where, and how.
- A wireframe represents the simple structure of a design. It can describe the functionality of a product and what happens for example what happens when you click a certain button. Decisions on what content to put where in the design are often made using wireframes.
- Mock-ups provide medium-fidelity representations of a design and design elements such as colours, fonts, texts, images and logos are used for the mock-up.

- Paper prototypes of a user interface can be created with pencils, paper, paint, sticky notes, cards, and coloured paint or paper. In paper prototyping, users can perform realistic tasks by interacting with a paper representation of the interface.
- Digital prototyping mainly has the same principle of paper prototyping, and the interface is represented in digital form in order to give the user an idea about the final outcome and enables the designer to make the prototype clickable and interactive [5].

In this thesis we concentrate on digital prototypes, however working with non-digital prototypes is also a good option, since it enforces teamwork and has been so far a cheap option to get early feedback from users [2].

2.2. User Experience and Artificial Intelligence

The User Interface (UI) encompasses the visual components, layouts, and graphical elements that enable user interaction with a digital system. It includes design features such as buttons, icons, and overall visual aesthetics that contribute to the usability and user experience of the product [23].

User experience (UX) is a consequence of a user's internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.) and the context or the environment within which the interaction occurs (e.g. organizational/social setting, meaningfulness of the activity, voluntariness of use, etc.) [7]. User experience encompasses all aspects of the end-user's interaction with the company, its services, and its products. It is important to distinguish between a well-designed User Interface (UI) and a positive User Experience (UX), as an effective UI constitutes only one component of a comprehensive and high-quality [24].

Artificial Intelligence (AI) refers to the simulation of human intelligence by a system or a machine. The goal of AI is to develop a machine that can think like humans and mimic human behaviours, including perceiving, reasoning, learning, planning, predicting, and so on [19].

Generative AI is rapidly emerging transformative innovation that enables machines to autonomously generate novel content—such as text, images, audio, or code—that mimics human-like creativity and expression, by learning from large datasets [27]. Generative AI contain a discriminator or transformer model trained on a corpus or dataset that is able to map the input information into a latent high dimensional space and a generator model, that is able to generate an stochastic behaviour creating novel content in every new trial even from the same prompts as an input [6]. The key aspect about generative models is that their architecture and the data amount that they have been fed is huge and that sets extremely high limits for computing power. For example, it is possible now to estimate the parameters of the model by feeding it the contents of the whole Wikipedia, Github, social networks, Google images and many more [6]. Generative models can learn the complex structures found in language or images. When these models are combined with systems that can translate the underlying meaning into different media formats—such as text, sound, or video—it becomes possible to transform input in one form (for example, written language) into output in another form (such as a video). Applications of this technology are endless, in the sense that models can be trained to generate genuine different multimedia formats such as video, audio or text from different multimedia input formats as well as from text [6]. Generative Adversarial Networks (GANs) are widely used as the primary approach for text based image generation in text to image models. They allow for efficient sampling of high resolution images with good perceptual quality, but are difficult to optimize and struggle to capture the full data distribution [14]. Generative AI architectures such as generative adversarial networks (GANs) and diffusion models are designed for generating realistic synthetic data, such as images, audio, text, and videos. Diffusion models have emerged as a dominant framework, especially when integrated with text encoders such as CLIP and T5, enabling precise text-conditioned image generation.

Large Language Models (LLMs) represent a breakthrough in the field of natural language processing (NLP) and artificial intelligence (AI). They achieved a significant commercial breakthrough following the release of OpenAI's ChatGPT in November 2022 [9]. These models, powered by deep learning algorithms, have shown remarkable capabilities in understanding, generating, and interpreting human-like text, fundamentally altering the landscape of computational linguistics and AI-driven applications [9]. LLMs are a special class of pretrained language models (PLMs) obtained by scaling model size, pretraining corpus and computation. LLMs, because of their large size and pretraining on large volumes of text data, exhibit special abilities which allow them to achieve remarkable performances without any task-specific training in many of the natural language processing tasks. It started with OpenAI’s GPT-3 model, and the popularity has increased exponentially after the introduction of interfaces like ChatGPT. The GPT4 model also accepts images besides text as input [8]. The next section outlines the related work done in context of this thesis.

3. Related work

Since the landscape is changing fast and new products with features appear it is important to keep in mind the time of publications in order to keep the analysis and conclusions relevant. November 2022, when the first commercial version of Chat GPT by Open AI was released, has been taken as a starting point for selecting research papers to review in this section.

A pertinent study was undertaken to develop a fitness mobile application prototype designed for next-generation users, in which Generative AI was employed as a co-creator throughout the Design Thinking process, including the prototyping stage. Kamnerddee, Putjorn and Intarasirisawat conducted and experiment where two low-fidelity wireframe versions were created: one manually designed by human UX/UI professionals using Figma[2], and the other generated by AI using Visily. These prototypes were assessed through A/B testing and qualitative interviews. The findings reveal that while AI-generated designs demonstrate strengths in creativity and efficiency, they often fall short in capturing the subtle nuances essential for human-centered design—for instance, inconsistencies in font usage across screens. Nonetheless, the authors of the paper concluded that AI-based prototyping tools show considerable potential, as they enhance the creative process by offering innovative design solutions that may not be immediately evident to human designers. An additional benefit is the acceleration of the prototyping process when AI tools are employed, making them valuable assets in design workflows [21].

In another paper [22] the use of generative AI in prototyping the authors found that it can serve as a catalyst for the creative process, as it fosters innovation and promotes experimentation by transforming failure into a productive pathway that often leads to improved ideas. Prior to engaging with generative AI, it is essential to clearly define the intended goals and identify the target audience for the outcomes. The process should be approached in reverse: begin by selecting appropriate media and AI tools, then proceed through iterative development informed by testing and feedback. Generative AI can be employed in the creation of mock-ups, visual inspirations, characters, scenes, recipes, or code. It also supports the exploration of conceptual variations and can significantly enhance collaborative efforts, such as paper prototyping or storyboarding. A critical factor for success is iteration—continuous experimentation and refinement help to clarify both the problem and its potential solutions [22].

[2] https://www.figma.com/

A systematic literature review[16] examines how artificial intelligence (AI) is utilized across various stages of the UX design process. These stages include understanding the context of use, uncovering user requirements, supporting solution design, evaluating design outcomes, and assisting in the development of solutions. Within the context of prototyping, the review emphasizes the potential of AI to facilitate transitions between low, mid, and high-fidelity prototyping. This capability enables automated and rapid iterations between stages, streamlining the process and allowing, for instance, the visualization of a user interface (UI) based on an initial sketch. Given the rapid advancements in AI, the authors suggest that future seminal publications will emerge, potentially reshaping the interpretation and application of AI within the UX design domain [16].

Another paper [15] was undertaken to explore the applications of large language models (LLMs) within the design process, specifically examining their potential to replace or augment human input. Based on the findings, four principal guidelines were established to define the effective integration of AI tools within the design workflow.

1. Be transparent and honest about where AI tools were used and provide information on how they augmented HCD or fell short.
2. Be aware of limitations and biases. Actively mitigate them in prompts and by checking the examples.
3. Involving the user is not an end in itself. Hence, if we can get information from LLMs of the same quality, we should not waste human resources.
4. Utilize LLMs to create functional prototypes to improve what users experience [15].

Another research paper[17] about AI-enhanced design thinking was published in April 2024. At the heart of its exploration was the belief that AI, when thoughtfully integrated into design thinking, not only complements but also significantly enhances human capabilities in creativity and problem-solving. However, it requires a delicate balance between leveraging technological advancements and maintaining the human-centric essence of design thinking. In order to ensure that AI acts as an enabler rather than a disruptor of human creativity and empathy the author outlined guiding principles that can navigate these challenges [17]. Since the focus of this thesis is evaluating prototyping tools, deep dive into the frameworks' challenges and mitigation actions will not be made. In the same paper it was also tested and provided evidence about the results of different AI enabled tools in various parts of the design process. Also several other use cases with prompts and results which illustrate the practical capabilities of AI in both problem identification as well as in the solution development phase were provided. A use case about prototyping was provided as well, where the author used Uizard application to create visual wireframes based on text prompt. According to the author the wireframes' design looked modern [17].

Bilgram and Laarmann[1] used ChatGPT with GPT4 in the early phases of innovation, i.e., exploration, ideation, and prototyping while trying to create an app for the automotive industry. Although they provided use cases and analysis for exploration and ideation as well, their focus was on AI augmented prototyping of digital products [1]. They found out that LLMs via agents like GPT can operate as a text-to-code generator empowering users to embrace early prototyping without writing code themselves. Non-technical users simply describe the digital product (e.g., a website) in natural language, which the LLM then turns into programming code (e.g., HTML code). This puts nontechnical users in a position where they are able to perform the exploration, ideation and early look- and feel-like prototypes that can be tested by user groups. Although they were not able to create a prototype directly from a text prompt they made progress after prompting *"Guide me through the process of creating the code for a simple prototype that I can open with my browser"* and got step by step instructions on how to create a prototype. The instructions did not require advanced technical knowledge but they assumed that the reader was familiar with text editors, such as Notepad or VS Code.

In their experiment the authors did not not write prompts using professional design language and did not act as UX specialists, but rather embodied as users with basic knowledge about the design process. Their generated website initially appeared to be very basic, but it improved after several communications with Chat GPT regarding enhancements of new features, design elements and navigation. The authors concluded in their work that with their prototype website users could evaluate general feasibility and functionality but not design and development nor test product acceptance in user interviews due to a limited visual quality. Nonetheless, the authors anticipated rapid progress in the development of new, specialised AI design tools that can create prototypes with both conceptual depth and high-quality design and render obsolete the better part of manual digital prototyping in early phases [1].

Another study about availability of AI enabled prototyping tools, focusing on mobile app development, was conducted in November 2023 by Stephan Bohm and Stefan Graser [3]. They carried out an experiment by trying to create a prototype for a mobile app with the help of AI tools. As a result of their work, the researchers identified and evaluated three forms of AI support in prototyping. First, they explored AI-based prototyping plug-ins, highlighting the

Figma plug-in Wireframe Designer[3], which utilized the ChatGPT-3.5 API to generate mock-ups from text prompts, transform them into a machine-readable format, and visualize them in Figma. Second, they examined prototyping support by general generative AI, using ChatGPT to develop a simple functional augmented reality (AR) prototype without requiring programming expertise. Third, they assessed integrated AI-based prototyping solutions that employ generative AI to convert textual input directly into visual prototypes, often with embedded additional AI functionalities. They concluded that by the end of 2023, while AI-supported tools could produce initial mobile app designs and aid idea generation—particularly for standard applications and small companies—the tools remained in early (beta) stages and were not yet capable of replacing professional designers[3].

Related literature is summarized in Table 2.

Table 2. Related literature.

Title	Author(s)	Publication year
AI-Driven Design Thinking: A Comparative Study of Human-Created and AI-Generated UI Prototypes for Mobile Applications	Kamnerddee, C., Putjorn, P., Intarasirisawat, J	2024
Prototyping with Generative AI: Augmenting Creative Workflows with Generative AI	Pennefather, P. P.	2023
Simulating the Human in HCD with ChatGPT Redesigning Interaction Design with AI	Schmidt, A., Elagroudy, P., Draxler, F., Kreuter, F., & Welsch, R	2024
AI-enhanced Design Thinking	Tschepe, S	2024
Accelerating Innovation With Generative AI: AI-Augmented Digital Prototyping and Innovation Methods	Bilgram, V., Laarmann, F.	2023
AI-based Mobile App Prototyping Status Quo, Perspectives and Preliminary Insights from Experimental Case Studies	Böhm, S., Graser, S.	2023

[3] https://www.figma.com/community/plugin/1228969298040149016/wireframe-designer

Artificial intelligence (AI) for user experience (UX) design: a systematic literature review and future research agenda	Stige, Å., Zamani, E. D., Mikalef, P., Zhu, Y.	2023

The related literature enables the author to identify the context and identify the available applications necessary for conducting the research and addressing RQ1. In the next chapter methodology will be outlined.

4. Methodology

This chapter provides an overview of the research process employed in the study, with a detailed outline of the specific steps undertaken throughout the process, including the participant evaluation with participants. The chapter is divided into 4 subsections, addressing the activities performed. This thesis is guided by a case study research method which is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident [20].

4.1. Overview of AI-enabled prototyping tools

The research process started with data collection phase what involved identifying applications that meet specific criteria for creating digital prototypes with AI assistance. The criteria are:

- The application must support the creation of prototypes.
- The application must include integrated AI functionalities that facilitate the prototyping process.

The activities undertaken to gather relevant applications are accessible in Appendix V. As a result of the data collection process, 20 applications were identified. For further application screening specific criteria were established to screen the initial set of results. The criteria for selection were as follows:

- The applications were accessible via web browsers as of the testing date (October 2024).
- The applications' web pages indicated the availability of AI functionality, either in beta, free, or paid versions.
- The application included a functional input field for entering text prompts, which was operational during testing.

As a result of the application screening seven applications out of the initial twenty met the criteria. Next, seven selected applications from the previous step were tested and evaluated against new criteria. A heuristic evaluation method [25] was employed, involving visual inspection of the user interfaces to assess functionality. The applications' testing criteria were:

- Selected application enables to communicate with application using text prompts and by that enables:
    - to create and modify prototype's pages

    - to edit design using text prompts
    - to create and edit interactions between prototype pages and elements
- The selected application can produce high-fidelity prototypes that meet the criteria of heuristic evaluation.

After testing seven applications, two were identified that met the criteria: Uizard [4] and Visily. Since both are commercial applications, evaluating their full versions involves monetary costs. Visily was chosen due to its lower price, despite they offered similar functionality.

4.2. Prototype Development

Requirements were established to initiate the prototype development in Visily. They were defined solely by the author based on observation of the functionality of similar available applications. To minimize potential bias in defining the requirements, they were formulated prior to installing the full version of the Visily application. The prompt for creating a prototype is accessible in Appendix VI. The detailed prototype creation process is outlined in Appendix VII.

4.3. Participant Evaluation

A remote participant evaluation was conducted between 13–14 March 2025 to identify usability issues within the prototype. Microsoft Teams was utilized for this purpose, employing its video call and screen-sharing functionalities. To facilitate subsequent analysis of the approximately 15–20 minute sessions, the built-in recording and transcription features of Teams were also employed. To conduct participant evaluation [13] five users were recruited. The selection of users covered three men and two women between ages 20-50 with digital literacy. Prior to conducting the evaluation, three questions accessible in the Appendix I, were surveyed regarding their digital experience to assess their familiarity with digital applications. Participants were asked to rate their level of comfort in using new digital applications without instructions on a five-point scale. To protect the privacy of participants, each participant was assigned a code in the form of [P1], [P2]…[P5]. Three specific tasks were asked to perform during remote guided testing sessions. Detailed descriptions of steps of tasks were not provided to participants during testing sessions.

[4] https://uizard.io/

During the task completion users were asked to think out loud. The success of tasks is evaluated based on predefined success criteria using a three-tier metric:

- Complete Success – The user successfully completes the task without encountering any difficulties.
- Partial Success – The user completes the task but experiences some difficulties.
- Failure – The user is unable to complete the task.

Tasks with success criteria are described in Table 3.

Table 3. Participant evaluation tasks and success criteria.

Tasks	Success criteria
1. Rent a bike from bike station next to the bridge	The participant was able to rent the bike.
2. Find a rental station where you rented a bike on September 5, 2023	The participant was able to find the rental station’s name.
3. Find out the maximum rental period of bike rental service	The participant was able to identify the rental period.

After completing the tasks, participants were asked to answer questions about overall experience, pain points, ease of use, expectations and improvement ideas in a follow-up interview. The questions are accessible in Appendix VIII.

4.4. Iterative Design Based Upon Feedback

Based on participant feedback new requirements were established and changes were made into a prototype. Appendix III presents a list of prompts used in implementing the requirements for improvement ideas. It is important to note that not all suggestions derived from participant evaluation were converted into formal requirements. This is due to the necessity of extensive backend development which was not feasible in Visily.

In the next chapter findings to research questions are presented.

5. Findings

This section outlines the findings of research questions. The first subsection 5.1. Current Landscape of AI-enabled Prototyping Tools answers RQ1: *What is the current landscape of AI-enabled prototyping tools?”*. Next subsection 5.2. Usability of AI-generated Prototype presents the answer to RQ2 *“How do users assess the usability of an AI generated prototype?”* and last subsection 5.3 Iterative Improvement of AI-generated Prototype focuses to answer RQ3: *“How can an AI-generated prototype be iteratively improved based on user feedback?”*

5.1. Current Landscape of AI-enabled Prototyping Tools

This subsection presents the findings to RQ1: “*What is the current landscape of AI-enabled prototyping tools?”*

The research commenced with data collection aimed at identifying AI-supported prototyping tools. Through web searches and ChatGPT queries, 20 applications were initially identified. However, only 7 of these applications were accessible via web browsers, demonstrated some degree of AI functionality, and included an operational input field for prompt submission. The list of these applications is presented in Table 4. Since only Uizard and Visily can be considered as dedicated tools for prototyping, the suggested primary use of other applications is outlined in the table.

Table 4. Analyzed applications.

Name of Application	Address	Primary use
Visily	https://www.visily.ai	Prototyping tool
Uizard	https://uizard.io	Prototyping tool
Galilieo	https://www.usegalileo.ai	Design creation
Relume	relume.io	Wireframes’ creation tool
Jeda	jeda.ai	Ideation tool
Dora	dora.run/	Web design tool
Banani	https://www.banani.co/	Design creation

Others initially identified applications had either no AI functionality at all, unusable low performance, they were discontinued, or it was evident that they were not intended for prototyping but rather for other tasks. Furthermore, besides applications there were also some AI supported plugins identified, like Text-to-Design, for collaboration tools such as Figma which can be related to prototyping. However, there was no evidence that any individual plugin could generate a complete prototype, rather they were for specific design tasks.

Uizard and Visily were applications that fulfilled the criteria outlined in Step 3 of the methodology section and were capable of generating functional prototypes. It was evident that these applications were also the most advanced in terms of user experience and overall functionality. Visily will be discussed in greater detail in the subsequent subsections, particularly in relation to the findings of Research Questions 2 and 3. Uizard, an application with capabilities similar to those of Visily, successfully generated a navigation flow complete with interactions and design elements based on a textual prompt. Although only the free version of Uizard was used—limiting the extent of experimentation with the tool—the initial navigation flow it produced was consistent with the prompt provided. The prompt used was: *Create a prototype for a community bike sharing app with 5 screens, displaying bikes available nearby, payment info, rental activity and instructions for unlocking the bike. Create interactions between elements. Clean and minimal design.* The same prompt was also utilized in testing other applications. Figure 1 illustrates the flow generated by Uizard. It is important to note that this output was produced on the first attempt, without any manual modifications or interventions.

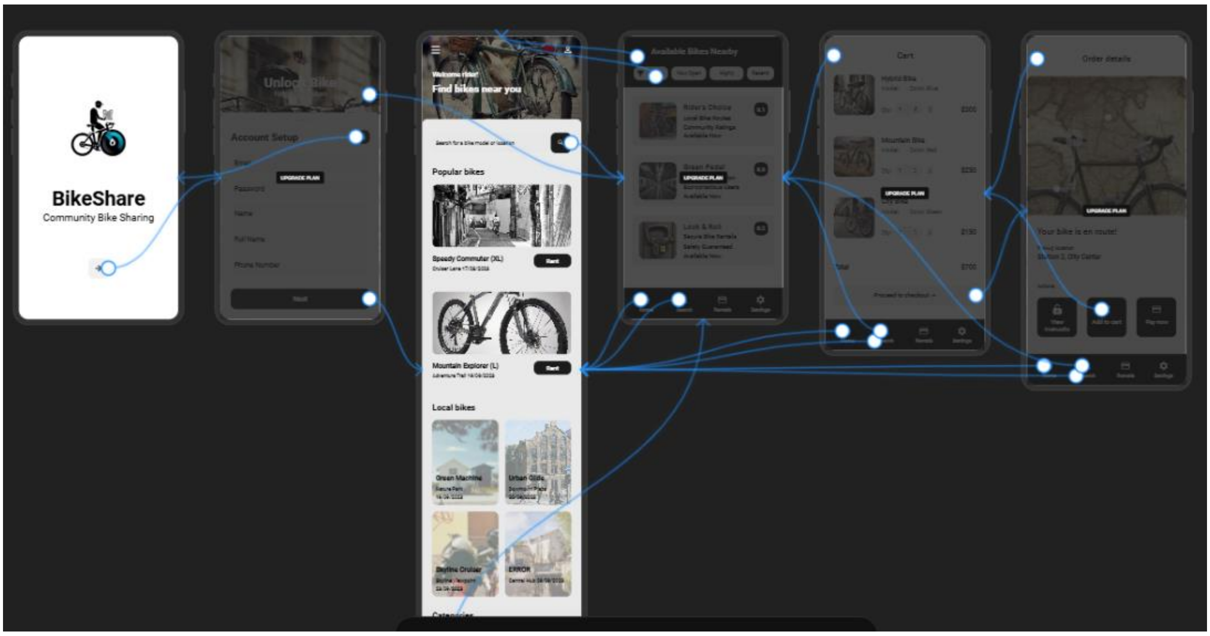


Figure 1. Illustration of Uizard

The other five applications demonstrated substantially lower output quality, and their primary focus was not on prototyping. Among them, Galileo[5] offered a positive user experience—particularly through its chatbot interface—yet the quality of its AI functionality in the trial version was limited. The AI failed to complete even a simple task, such as adding a navigation bar to a page. This limitation is illustrated in Figure 2. Nevertheless, the option to export created pages to Figma provided a valuable workaround, enabling further development within the Figma environment.

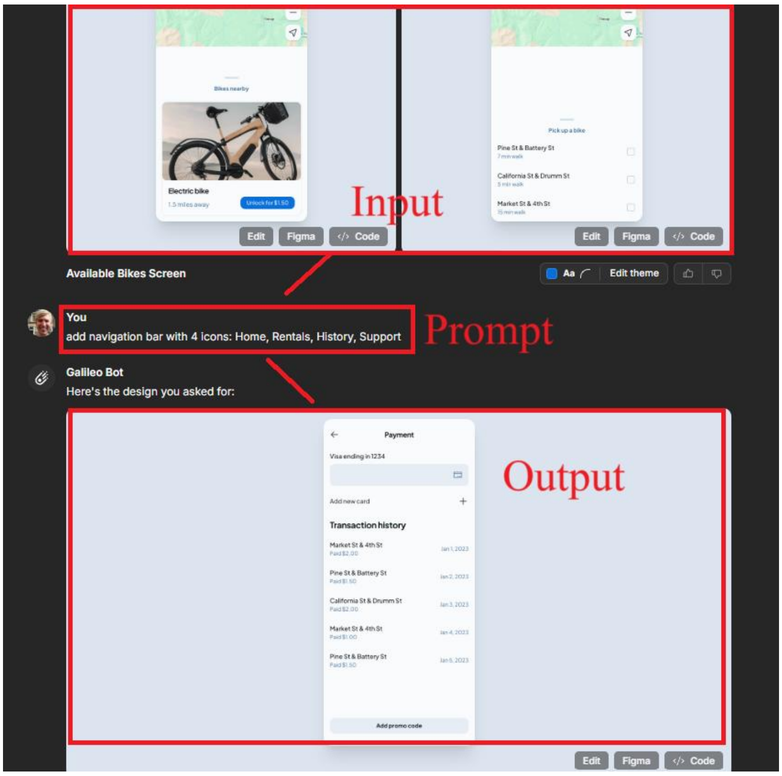


Figure 2. Illustration of Galileo's AI limited outcome

Relume[6] was an application that enabled the creation of wireframes using AI. However, its free trial was restricted to a single iteration. Jeda[7] was a platform that integrated various AI-enabled functionalities, among which wireframe creation was one option. Dora[8] could generate a visually colourful design; however, it produced responsive web pages rather than the mobile application specified in the prompt. Additionally, the free version did not support the creation of multiple pages simultaneously, making it impossible to establish a navigational flow

5 https://www.usegalileo.ai/
6 https://www.relume.io/
7 https://www.jeda.ai/
8 https://www.dora.run/ai

between screens. Compared to the other applications tested, the beta version of Dora was notably slow.

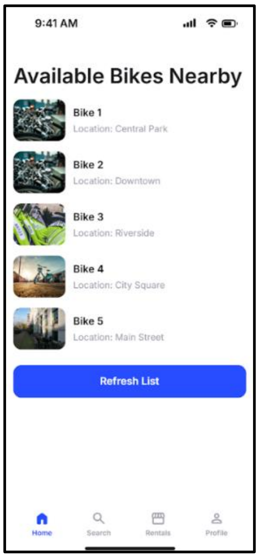


Figure 3. Example of Banani

Banani[9] was an application that facilitated the creation of wireframes with the assistance of AI. While editing pages via text prompts was effective, the platform lacked support for establishing flow and interactions between pages. Figure 8 illustrates the output generated by Banani.

The primary finding of this section indicates that AI-enabled prototyping applications are still in the early stages of development, with only two tools—Visily and Uizard—demonstrating a sufficient level of maturity to be effectively employed in prototyping tasks.

5.2. Participant Evaluation of AI-generated Prototype

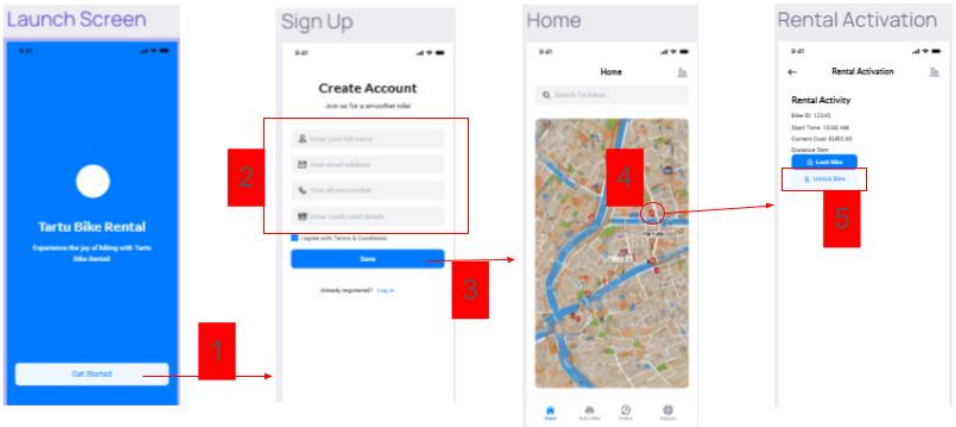
In this section the findings RQ2: "*How do users assess the usability of an AI generated prototype?*" will be covered. The discussion is substantiated by participant evaluations regarding the deployment of the prototype developed using Visily AI.

At the beginning of evaluation participants were asked to rate their level of comfort in using new digital applications without instructions on a five-point scale. Their responses ranged from scale values 3 to 5.

The tasks evaluated by participants are outlined in Section 4.3, Participant Evaluation.

[9] https://www.banani.co

TASK 1: Rent a bike from bike station next to the bridge **SUCCESS CRITERIAS:** The user was able to rent the bike.



STEPS TO PERFORM: 1. Navigate from Launch screen to Sign-up screen by clicking a button Get started 2. Conduct registration by filling in name, e-mail address, phone number and credit card number on the Sign-up screen 3. By clicking Save button on Sign up page navigate from Sign-up screen to Home screen 4. On the map on Home screen click on the icon of rental station station what is next to the bridge 5. On Rental Activation screen click on button Unlock Bike.

Figure 4. The description and steps to perform of task 1

Task 1 asked participants to "Rent a bike from the bike station next to the bridge." Different bike stations were displayed on a city map on the home screen. All participants easily navigated from Launch screen to Sign Up screen (step 1) and successfully completed the registration process (step 2). They found it relatively easy to navigate. However, participant [P1] raised concerns about the security of entering credit card details by saying: *"Already asking about the credit card details? To be honest, I wouldn't actually be sure that I would actually enter my credit card information here when I'm creating this account. It's a bit fishy because I have absolutely no idea what is happening after…"*. To address that concern during the test, participants were instructed to enter fictitious credit card information. On Figure 4 task 1 is described with success criteria and detailed steps to perform. After saving the registration, the prototype displayed a Home page featuring a map with rental stations (step 3). A common challenge among participants was the low quality of the map and difficulty in locating the correct rental station (step 4). Figure 5 illustrates the challenges of finding the right rental station next to the bridge.

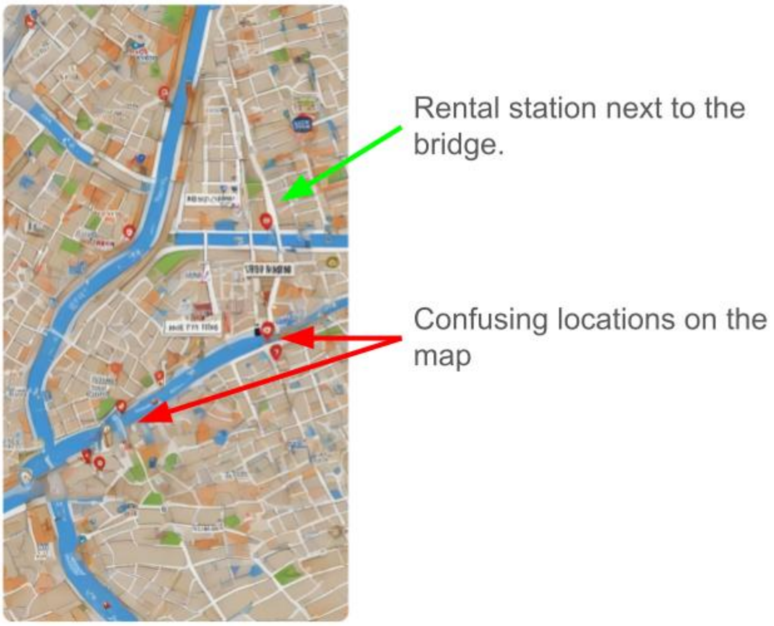


Figure 5. Visual representation of challenges displayed on the map within the Home screen

Additionally, two participants initially assumed that the prototype included real-time geolocation functionality on the map, which was not the case. Furthermore, since Visily was unable to generate an actual map of Tartu, the use of a generic city map led to additional confusion, as all participants were familiar with the geography of Tartu. One user addresses her confusion by saying *"...so much, like stuff going on here and there. So I don't understand all the streets, street names…"*. Figure 6 illustrates the confusion of users during the participants evaluation. Despite these challenges, all participants were ultimately able to identify the correct rental station on the map and proceed to the rental activation page, where they selected the "Unlock Bike" button (step 5).

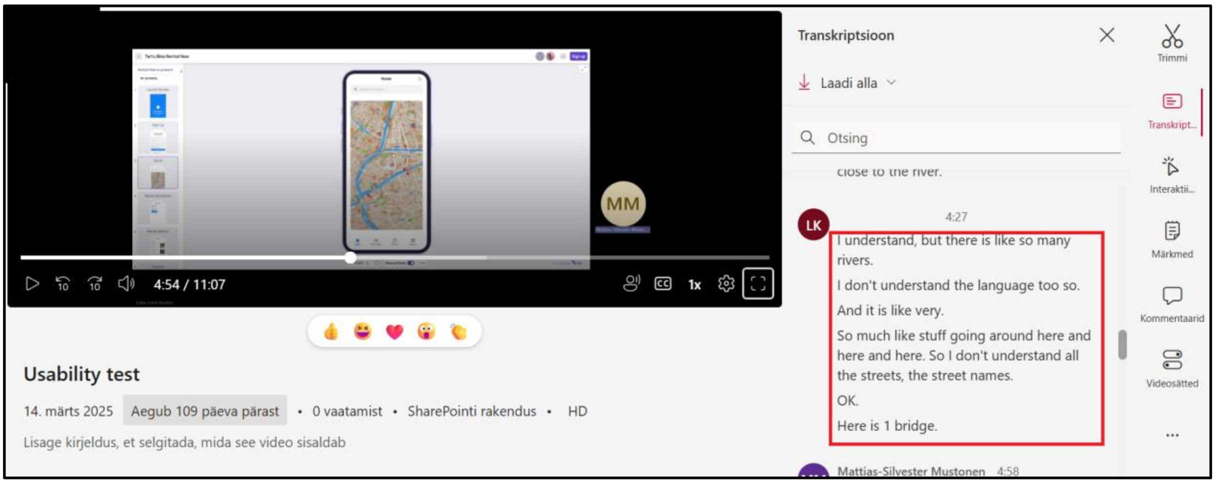


Figure 6. Snapshot from participant evaluation highlighting the confusion about the map's geography

During step #5, participant [P1] provided feedback regarding usability issues on the rental activation page by saying: *"...the sequence of these buttons doesn't make sense to me, because first I would want to unlock the bike"* and further suggesting modifications to the sequence and highlighting of buttons for improved clarity.

Additionally, all participants found the information displayed about the ongoing rental period on the rental activation page confusing, as they mistakenly assumed that the displayed distance information referred to the rental station rather than the rental session itself. Figure 7 illustrates the issues on the rental activation page.

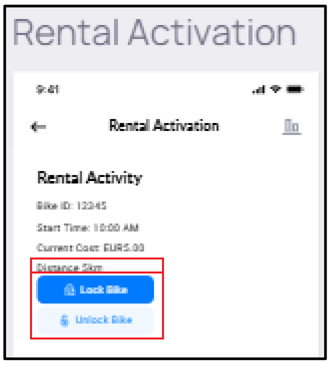


Figure 7. Issues on the rental activation page

During the follow-up session users provided additional feedback. Four users appreciated the familiar behaviour of the interface and positively highlighted the organization of its elements. One participant [P3] indicated that "...*the prototype works pretty well, it is in the standard module, this is good and universal for customers"*. However, another participant [P2] was critical of the overall experience due to difficulties encountered while navigating the map. When participant [P2] was asked to indicate one thing to improve, she answered: "*The map. 100%"*.

All participants agreed that the prototype's map functionality was insufficient to effectively illustrate the bike rental process. Specifically, they noted the absence of geolocation features for both participants and available bikes, as well as missing details such as distance, price, and availability on the rental station pins. Participant [P4] also suggested that the overall design could be more visually engaging. Additionally, participant [P5] expressed confusion regarding the function of the search bar, questioning what it would search for if it were operational. Although other participants did not explicitly raise this concern, observations during participant evaluation indicated that multiple participants attempted to use the search bar during Task 1, reinforcing the importance of addressing this issue.

Based on a three-tier evaluation scale, Task 1 was classified as a **partial success** which indicates, based on defined methodology, that the participant completes the task but experiences some difficulties. While all participants successfully completed the task, they encountered notable difficulties during the process.

Task 2 required participants to complete the action: *"Find the rental station where you rented a bike on September 5, 2023."* All participants began this task from the Rental Activation page, which was the final screen of Task 1. The steps to perform are described in Figure 8.

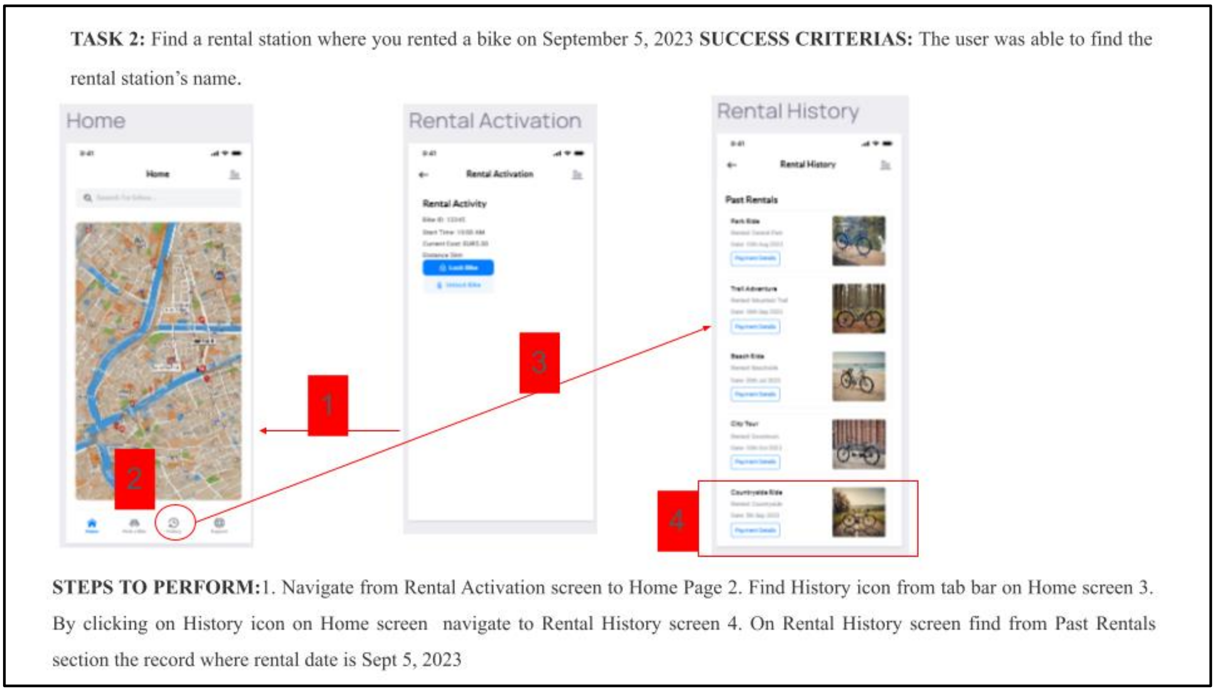


Figure 8. The description and steps to perform of task 2

To locate the required information, participants needed to navigate back to the Home screen and then use the tab bar to access the History icon, which directed them to the Rental History page. The requested rental information was located at the end of the list of past rentals.

This task was evaluated as **complete success** which indicates, based on methodology, that the participant successfully completes the task without encountering any difficulties. *"History was* *the easiest task to find"*, was a quote provided by one participant [P5]. All participants were able to efficiently locate the rental station's name, which was a success criteria for task #2. During the process, one participant [P1] noted a usability issue by saying that *"buttons down there are crawlable, and they are hidden",* indicating to the partially hidden tab bar on the Home screen what was somewhat difficult to identify the correct icon. Figure 9 presents a screenshot from the interview in which the issue was discussed.

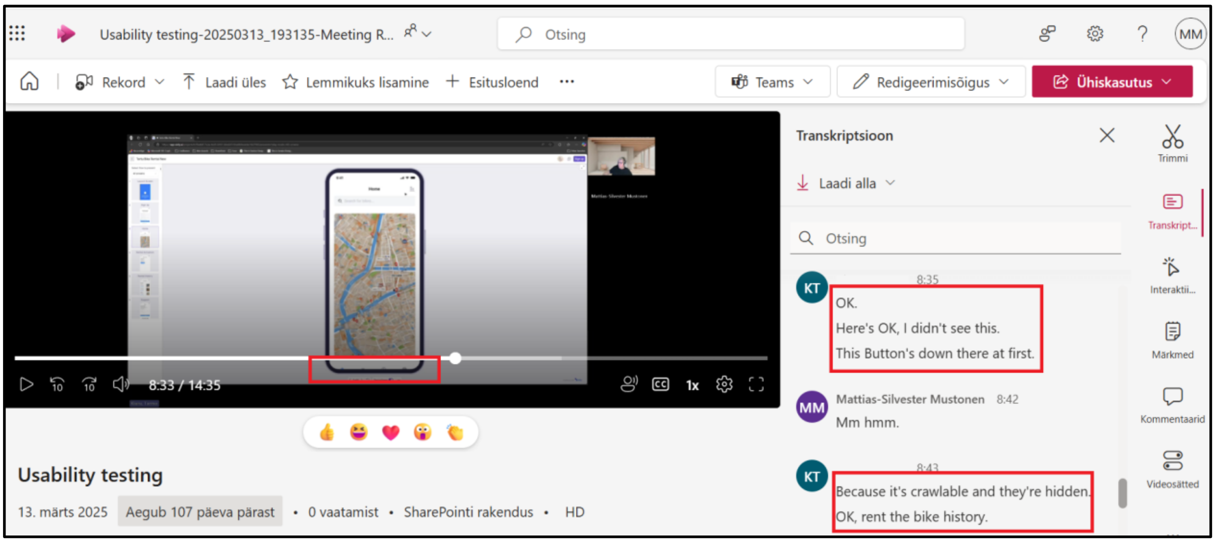


Figure 9. Screenshot from the participant evaluation interview in which the issue of the partially hidden tab bar was identified

Task 3 asked to *"Find out the maximum rental period of bike rental service"*. This information was in the Frequently Asked Questions section on the support page, which could be accessed through the Support icon on the tab bar. During testing, the navigation flow began from the Rental History screen, which was the final screen visited during the execution of Task 2. The task description and corresponding steps are illustrated in Figure 10.

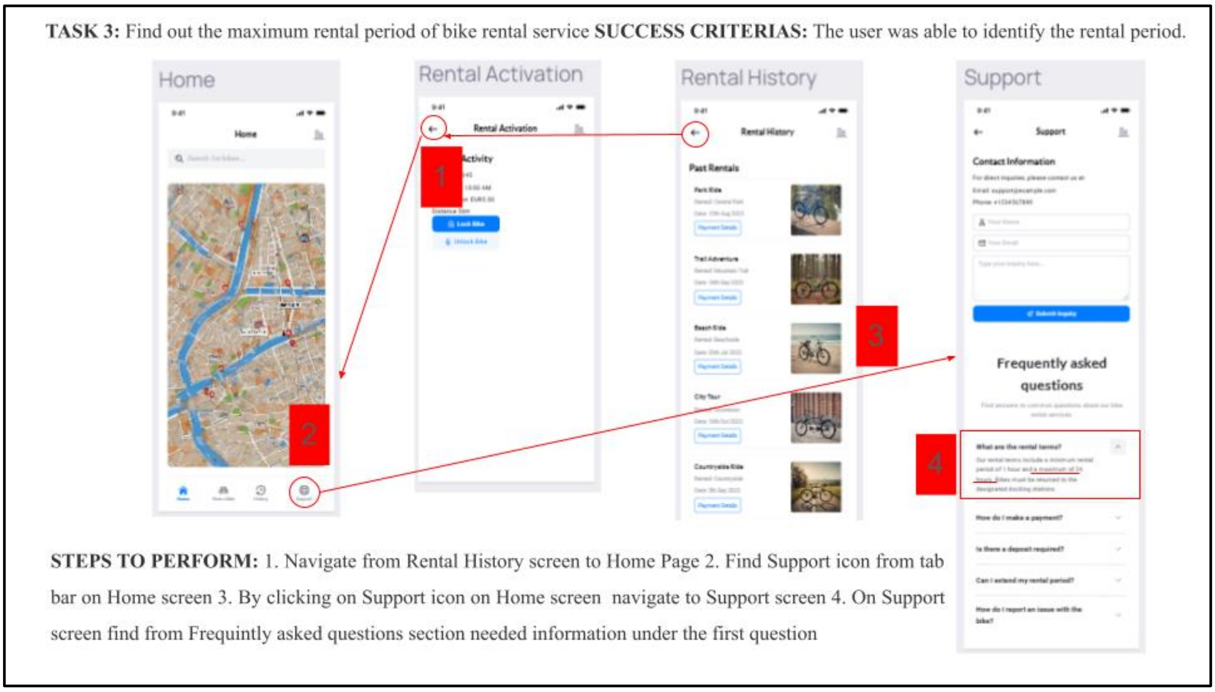


Figure 10. The description and steps to perform of task 3

Four participants intuitively navigated to the support page and quickly located the relevant section. However, participant [P1] initially assumed that the rental terms would be displayed during the rental activation process and therefore did not search for the information on the support page. As a result, it took for a participant a considerable amount of time to realize that the necessary details were in the support section. With some guidance, the participant was eventually able to complete the task. Regarding the evaluation, Task 3 was evaluated against established criteria as **partial success,** which means that the participant completes the task but experiences some difficulties based. Table 5 summarises the results of performed participant evaluation.

Table 5. The results of participant evaluation.

Task	Result based on three-tier metric
1. Rent a bike from bike station next to the bridge	Partial success
2. Find a rental station where you rented a bike on September 5, 2023	Complete success
3. Find out the maximum rental period of bike rental service	Partial success

The primary finding of this section is that the prototype created using Visily was sufficiently developed to support participant evaluation. As none of the participants failed any of the assigned tasks, this outcome provides evidence that AI-generated prototypes have the potential to serve as viable alternatives to manually created prototypes. Moreover, such tools may enable the facilitation of basic evaluation sessions even by individuals without professional UX expertise.

In the following section, the identified improvement ideas from participant evaluation will be translated into specific requirements, and a new, improved prototype will be developed accordingly.

5.3. Iterative Improvement of AI-generated Prototype

In this section, the requirements for prototype improvement will be defined, and the process of implementation within the Visily application will be described. This allows an answer to RQ3: *"How can an AI-generated prototype be iteratively improved based on user feedback?"*

Table 6 presents a list of improvement ideas, additional requirements, and their corresponding justifications.

Table 6. Additional requirements.

Captured improvement ideas	Determined as a new requirement (Yes/No)	Justifications
1. Enhance the map on the homepage to accurately reflect the geography of Tartu city.	Yes	The primary issue identified during participant evaluation was the poor quality of the map
2. Enable geolocation positioning on the map.	No	This requires a backend connection and further IT development.
3. Delete search bar from home page.	Yes	Enabling search bar's functionality necessitates a backend connection and additional IT development. To have better clarity on the prototype without working functionality, the deletion is appropriate.
4. Change the sequence of lock and unlock buttons on the rental activation page, so that the unlock button is in a higher position than the lock button.	Yes	Feedback from participant evaluation.

5. Switch the highlight colours between buttons on the rental activation page, so that the unlock button is highlighted (blue with white text) and the lock button not highlighted (light blue with blue text).	Yes	Feedback from participant evaluation.
6. Create a link from the rental activation page that navigates to the Support page.	Yes	Feedback from participant evaluation.

To amend specific page content, a selected section or the entire page must be uploaded to Visily's AI design container. During the implementation of requirements, it became evident that to achieve optimal results, the necessary changes should be divided into small tasks and implemented incrementally. This approach ensures that results are effectively anchored.

The process involves uploading pages or sections to the container, refining increments through prompts, and then transferring the updated content back to the board. Subsequently, the revised version is re-uploaded to the container for further refinements. In other words, complex prompts should be avoided, as AI may introduce unintended modifications alongside the expected changes.

It is important to highlight that Visily AI does not offer an undo option, reinforcing the necessity of careful incremental implementation.

Low map quality was identified as the most critical issue during participant evaluation. While attempting to implement Requirement #1: Enhance the map on the homepage to accurately reflect the geography of Tartu city, it became evident that Visily AI was unable to generate a city-specific map based on a given prompt. Furthermore, it was not possible to edit the existing image depicting the map.

For instance, when prompted with *"Create a street map for central Tartu city"*, the AI produced a distorted, stylistically inconsistent map that was entirely unusable. Figure 11 illustrates the AI-generated map, highlighting the limitations of the tool in accurately representing geographic details.

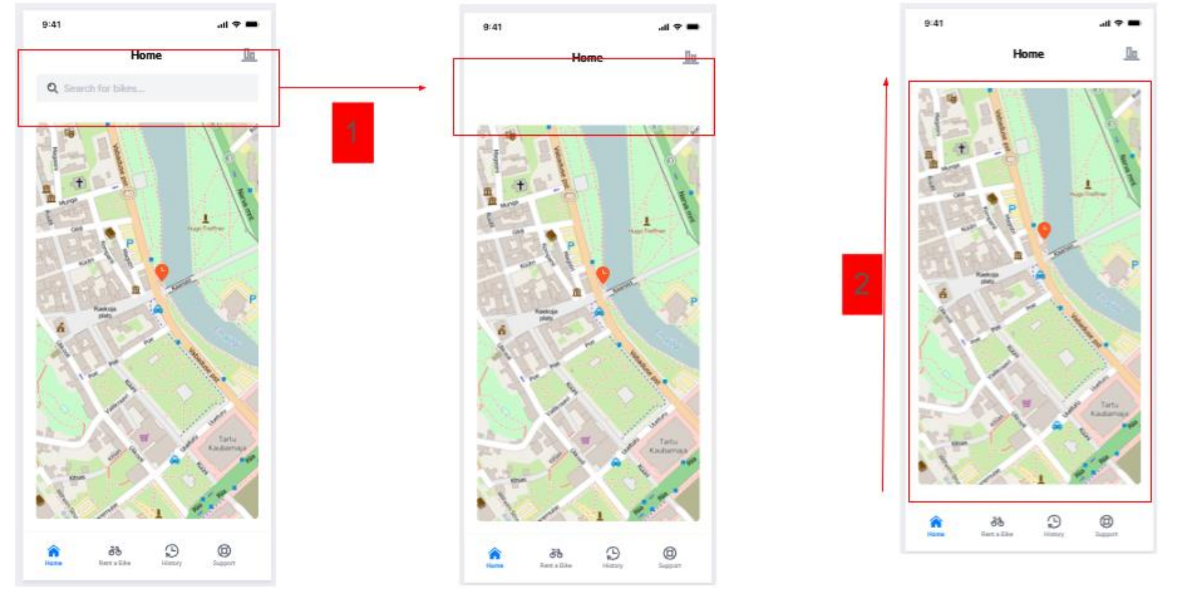
Figure 11. Map on central Tartu generated by Visly AI

To address this issue, an image of a map of central Tartu was manually uploaded to Visily. This process was intuitive and straightforward. Additionally, a location icon was manually added to the map to enhance its realism. However, it is important to note that both tasks were carried out manually, without utilizing Visily's AI functionality.

Since geolocation positioning requires backend development and real-time data integration, Requirement #2: Enable geolocation positioning on the map was excluded from the improvement process.

Requirement #3 was to remove the search bar from the homepage. By uploading the homepage with the new map into the AI container, the AI successfully eliminated the search bar using the prompt "Remove the search bar." Subsequently, another prompt was used to remove the empty space left between the screen title "Home" and the top of the map. The AI effectively adjusted the layout, eliminating the unnecessary space.

Figure 12 illustrates the evolution of the homepage as a result of these improvements.



Prompts used in Home page improvements: 1. Remove the search bar. 2. Remove the empty space left between the screen title "Home" and the top of the map

Figure 12. Evolution of the Home page improvements

Positive outcomes were achieved in implementing Requirement #4: Change the sequence of buttons and Requirement #6: Create a text link from the rental navigation to the support page. However, the navigation hotspot connecting the prototype to the support page via the text link had to be manually configured.

The evolution of the design, including the reordering of buttons, is illustrated in Appendix II. To maintain design consistency, efforts were made to ensure that the font colour of the terms matched that used on the homepage. It was not possible to fully implement Requirement #5: Switch the highlight colours of buttons. Despite multiple attempts and various prompts, the closest result achieved was one where the activation button was highlighted, but the background of the bike lock button became transparent. Visily appeared unable to properly contextualize the colour palette used on the page, often introducing random colours that did not align with the existing design. Figure 13 illustrates the result obtained in relation to the expected outcome of this requirement.

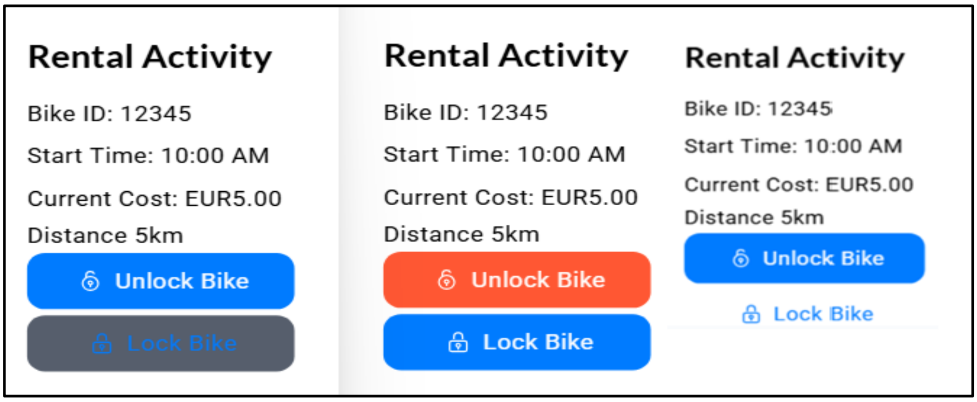


Figure 13. Various Results for Requirement #5 – Switching the Highlight Colors Between Buttons on the Rental Activation Page

All used prompts in implementing the requirements are in Appendix III. Table 7 summarizes the executed prototype's improvement ideas.

Table 7. Summary of executed prototype's improvement ideas.

Requirement	Result
Enhance the map on the homepage to accurately reflect the geography of Tartu city.	Not resolved
Delete search bar from home page.	Resolved successfully
4. Change the sequence of lock and unlock buttons on the rental activation page, so that the unlock button is in a higher position than the lock button.	Resolved successfully
5. Switch the highlight colours between buttons on the rental activation page, so that the unlock button is highlighted (blue with white text) and lock button not highlighted (light blue with blue text).	Not resolved
6. Create a link from the rental activation page that navigates to the Support page.	Resolved successfully

6. Discussion

This thesis explored the AI-enabled prototyping tools in user-centered design. The contributions of thesis were:

- Contribution to ongoing research of prototyping with AI
- Evaluation of AI created prototypes with participants
- The practical recommendations for creating prototypes with Visily

In addressing RQ1— "*What is the current landscape of AI-enabled prototyping tools?*"—a comprehensive investigation was conducted. The market for AI-powered applications that purport to support prototyping is evolving rapidly. Out of 20 initially identified applications, only seven progressed to the detailed screening phase. The remaining 13 were either in very early development stages, or exhibited performance levels too low to allow for effective testing.

The findings revealed that only two applications, *Uizard* and *Visily*, are specifically designed for prototyping with integrated generative AI support. This suggests that the landscape of AI-supported prototyping tools is still in its nascent stages.

Other applications identified during the data screening phase met the inclusion criteria in terms of generative AI capabilities and chatbot integration; however, their primary function was not to offer users comprehensive prototyping features. A common use case involved the generation of individual screen designs, which could then be exported to collaborative platforms such as Figma for further development—a feature that may nonetheless prove useful for designers.

In addition to standalone applications, plugins within collaboration platforms were also examined. However, no functioning plugins dedicated exclusively to prototyping were found.

In addressing research question 2— "*How do users assess the usability of an AI-generated prototype?*"—participant evaluation was carried out on a prototype created using Visily's AI functionality. The results indicated that, although the tool could generate a functional and visually coherent prototype based on user prompts, it lacked refinement in key usability aspects. Specifically, issues such as illogical button placement and a lack of context-aware design elements required manual adjustment. Participants also identified additional limitations, including low map realism and confusion regarding screen dimensions, which led to visibility issues with menu icons. It is important to note that the prototype's lack of backend integration contributed to some confusion during testing. This could likely have been mitigated through further refinement of the prototype prior to conducting user evaluations.

Research question 3— "*How can an AI-generated prototype be iteratively improved based on user feedback?"*—was addressed through practical experimentation. The most effective approach identified was to generate the entire prototype using a single, detailed prompt, and to iteratively refine it by re-submitting the prompt with necessary adjustments when the output did not meet expectations. However, as the editing feature in Visily is still in beta, making changes through prompts proved to be time-consuming; each query required 30 to 60 seconds to process, and multiple iterations were often necessary to achieve the desired outcome.

Visily also allows for manual adjustments, which emerged as the preferred method over relying on the chatbot interface for modifications. A significant challenge with AI-enabled editing of individual screens was the tool's limited contextual awareness across different pages. Because modifications occur within isolated containers, changes to one element frequently disrupt other elements on the same page. For instance, after several attempts, it was not possible to successfully change the highlight colour of buttons while keeping the other elements' design stable. Similarly, efforts to improve the visual quality of maps were unsuccessful due to minimal contextual integration. Additionally, initiating interactions between hotspots via the chatbot was not possible, although such connections could be easily established through manual editing.

Overall, current AI-enabled prototyping tools demonstrate considerable potential in generating initial high-fidelity prototypes. From a time-efficiency perspective, they can outperform manual prototyping methods. However, to maximize the benefits of AI capabilities, it is essential to prepare comprehensive prompts accompanied by design references. Fine-tuning and detailed adjustments, however, should be carried out manually to ensure accuracy and consistency.

6.1. Recommendations

Recommendation 1: When editing the generated prototype, it is advisable to use short, concise prompts. In Visily, modifications require uploading a page or element into a specific editing container. To minimize the risk of unintended changes—an issue that may arise with overly complex prompts—it is recommended to apply changes incrementally using brief input.

Recommendation 2: To provide Visily with appropriate context at the outset of prototype creation, it is recommended to run several initial iterations. Extracting themes from an existing webpage that serves as a reference can help guide the AI with expected design and improve output relevance and coherence.

Recommendation 3: For refining element placement and navigation flows, manual adjustments are recommended. Editing via AI can be time-consuming, as it often requires multiple iterations to achieve the desired outcome. Additionally, each iteration may take approximately 30–60 seconds to process, making manual refinement more time-efficient in certain cases.

6.2. Limitations

The limitation of the case study method is its limited generalizability, as findings are based on a specific context and may not apply broadly [26]. To mitigate this, we selected a case that is representative of typical conditions and documented the context in detail to support transferability. It should be noted that the scope of this thesis was limited to AI prototyping tools capable of generating high-fidelity prototypes. AI-based applications or plugins that might support other stages of the prototyping process were excluded from the analysis. Consequently, the study does not provide a comprehensive evaluation of the overall impact of AI technologies on the prototyping process. Furthermore, the number of AI supported prototyping applications was limited, since Visily was the only application used with its full-featured Pro version; all other applications were evaluated using free subscription plans, which likely limited access to certain advanced functionalities. To address the limited scope, the study focused on clearly defined criteria for tool selection and ensured consistent evaluation procedures across all included applications. While the exclusion of AI tools supporting earlier or adjacent stages of prototyping narrows the focus, it allowed for a deeper investigation of high-fidelity prototyping tools specifically. The use of free plans was acknowledged as a constraint; however, effort was made to assess core functionalities accessible to typical users, providing a realistic view of tool capabilities in common usage scenarios

7. Conclusion

Emerging AI technologies are poised to significantly influence both the process, and the outcomes of digital prototypes developed using digital prototyping tools. These tools facilitate close collaboration with users, enabling rapid feedback and iterative improvements to prototypes. The integration of AI capabilities into prototyping applications has the potential to enhance efficiency and expand design possibilities. However, it is essential to maintain realistic expectations regarding the capabilities and outcomes of AI-enabled prototyping tools—a topic that this thesis investigated.

Following the motivation presented in the introductory section, the study commenced in Chapter 2 with an overview of key concepts related to prototyping, the user-centred design process, and generative language models, providing essential contextual grounding. This was followed by a review of related literature, which offered an academic perspective through the analysis of relevant scholarly work.

In Chapter 3, the Methodology section outlined the research design, including the evaluation criteria and requirements applied to a prototype of the Tartu Bike Sharing App, developed using Visily—an AI-assisted prototyping application.

The Findings section synthesized the outcomes of the study's three research questions. The investigation began with RQ1: "*What is the current landscape of AI-enabled prototyping tools?"* Based on predetermined criteria—namely, the ability to interact via prompts and produce interactive prototypes—the analysis revealed that the landscape is still in its early stages. Only two tools, Uizard and Visily, met these criteria.

RQ2 explored: "*How do users assess the usability of an AI-generated prototype?"* A participant evaluation involving five participants was conducted on the Tartu Bike Sharing App prototype. Of the three tasks tested, one was completed successfully, while two were partially successful. The findings indicated that Visily was capable of generating a functional and visually coherent prototype, but exhibited shortcomings in several key usability aspects such as properly contextualising the colour palette used on the pages.

Insights from user testing led to the formulation of five new requirements, enabling the study to address RQ3: "*How can an AI-generated prototype be iteratively improved based on user feedback?"* Experimental results suggested that while prompt-based communication was effective for simpler tasks, the AI-enabled prototyping tool struggled with more complex concepts, such as applying a coherent colour palette or accurately rendering map details. Furthermore, the editing process proved time-consuming, with each prompt taking approximately 30 to 60 seconds to process.

Chapter 6 provided a detailed discussion of the findings in relation to the research questions.

In conclusion, the investigation titled AI-enabled prototyping tools in user-centered design proved to be a compelling and timely undertaking. Given the rapid pace at which AI technologies are evolving, similar studies conducted in the future could yield novel insights, making this a valuable area for continued scholarly exploration. Future research may benefit from a more detailed examination of the technical foundations of both existing and emerging AI tools, particularly regarding their output quality, performance, and environmental implications. Furthermore, the increasing integration of AI-generated multimedia content into design prototypes presents a range of challenges and opportunities—especially concerning technical feasibility and legal considerations, such as copyright and intellectual property issues.

Finally, the advancement of chatbot interfaces (e.g., ChatGPT) and their potential role in enhancing the prototyping process from a user experience perspective represents another promising avenue for future research.

Acknowledgments

A great deal of gratitude is given to the interviewees for finding the time to participate in the user research process and providing valuable insights.

References

[1] Bilgram, V, Laarmann, F. (2023, 06). *Accelerating Innovation With Generative AI: AI-Augmented Digital Prototyping and Innovation Methods* DOI: 10.1109/EMR.2023.3272799 Accessed at: 24.04.2025

[2] Ali, H. (2023, Oct). *Human-Computer Interaction: Enhancing User Experience in Interactive Systems* DOI:10.30572/2018/KJE/140403 Accessed at: 24.04.2025

[3] Böhm, S., & Graser, S. (2023, November). *AI-based Mobile App Prototyping Status Quo, Perspectives and Preliminary Insights from Experimental Case Studies*. Available at: https://personales.upv.es/thinkmind/dl/conferences/centric/centric_2023/centric_2023_1_40_38009.pdf Accessed at: 14.05.2025

[4] Brown, T. (2008, June). *Design Thinking*. Harward Business Review. Available at: https://designthinkingmeite.web.unc.edu/wp-content/uploads/sites/22337/2020/02/Tim-Brown-Design-Thinking.pdf Accessed at: 24.04.2025

[5] Burns, C. M. (2018, May). *Human-centred design. In book: eHealth Research, Theory and Development (pp.207-227), http://dx.doi.org/10.4324/9781315385907-10* Accessed at: 24.04.2025

[6] Gozalo-Brizuela, R., Garrido-Merchan, E. C. (2023, Jan). *ChatGPT is not all you need. A State of the Art Review of large Generative AI models*. Available at: https://doi.org/10.48550/arXiv.2301.04655 Accessed at: 24.04.2025

[7] Hassenzahl, M., Tractinsky, N. (2006, March-April). *User experience – a research agenda*. Behaviour & Information Technology, 25(2), 91 – 97. http://dx.doi.org/10.1080/01449290500330331 Accessed at: 24.04.2025

[8] Kalyan, K. S. (2024, March). *A survey of GPT-3 family large language models including ChatGPT and GPT-4.* Natural Language Processing Journal, 6. https://doi.org/10.1016/j.nlp.2023.100048 Accessed at: 24.04.2025

[9] Lecun, Y., Bengio, Y., Hinton, G. (2015). *Deep Learning.* Nature 521(7553):436-44. http://dx.doi.org/10.1038/nature14539 Accessed at: 24.04.2025

[10] Mackay, W. E. (2002). *Human-Computer Interaction: Development Process*. CRC Press. https://www.lri.fr/~mackay/pdffiles/Prototype.chapter.pdf Accessed at: 24.04.2025

[11] Mithun, A. M., Bakar, Z. A.,Yafooz, W. M. S. (2018). *Extended User Centered Design (UCD) Process in the Aspect of Human Computer Interaction.* International Conference on Smart Computing and Electronic Enterprise (ICSCEE) (pp. 1-6). Available at: Extended User Centered Design (UCD) Process in the Aspect of Human Computer Interaction | IEEE Conference Publication | IEEE Xplore Accessed at: 24.04.2025

[12] Pea, R. D. (1997, January). *User Centred System Design-New Perspectives on Human/Computer Interaction.* Journal of Educational Computing Research. Available at: https://telearn.hal.science/file/index/docid/190545/filename/A38_Pea_87e.pdf Accessed at: 24.04.2025

[13] Nielsen, J. *Usability Engineering* (1993), Morgan Kaufmann, San Francisco; ISBN 0-12-518406-9

[14] Rombach, R. (2021). *High-Resolution Image Synthesis with Latent Diffusion Models.* Available at: https://arxiv.org/pdf/2112.10752 Accessed at: 24.04.2025

[15] Schmidt, A., Elagroudy, P., Draxler, F., Kreuter, F., & Welsch, R. (2024, Jan-Feb). *Simulating the Human in HCD with ChatGPT Redesigning Interaction Design with AI*. Available at: https://dl.acm.org/doi/pdf/10.1145/3637436 Accessed at: 24.04.2025

[16] Stige, Å., Zamani, E. D., Mikalef, P., & Zhu, Y. (2023, Aug). *Artificial intelligence (AI) for user experience (UX) design: a systematic literature review and future research agenda.* Available at:

https://pdfs.semanticscholar.org/db91/4825e4ced8f1d1bddedff3d45ecf4c836e1f.pdf Accessed at: 24.04.2025

[17] Tschepe, S. (2024, April). *AI-enhanced Design Thinking*. Available at: DOI:10.13140/RG.2.2.35306.12484 Accessed at: 24.04.2025

[18] Walker, M., Takayama, L., Landay, J. A. (2002, Sept). *High-Fidelity or Low-Fidelity, Paper or Computer Choosing Attributes When Testing Web Prototypes.* Available at: DOI:10.1177/154193120204600513 Accessed at: 24.04.2025

[19] Xu, Y. (2021). *Artificial intelligence: A powerful paradigm for scientific research.* The Innovation, Volume 2(Issue 4). Available at: https://doi.org/10.1016/j.xinn.2021.100179 Accessed at: 24.04.2025

[20] Yin, R. K. (1984). *Case Study Research: Design and Methods.* Beverly Hills, CA: Sage Publications. Available at: DOI:10.3138/cjpe.30.1.108 Accessed at: 24.04.2025

[21] Kamnerddee, C., Putjorn, P., Intarasirisawat, J., (2024). *AI-Driven Design Thinking: A Comparative Study of Human-Created and AI-Generated UI Prototypes for Mobile Applications*, 2024 8th International Conference on Information Technology (InCIT), Chonburi, Thailand, 2024, pp. 237-242. Available at: AI-Driven Design Thinking: A Comparative Study of Human-Created and AI-Generated UI Prototypes for Mobile Applications | IEEE Conference Publication | IEEE Xplore Accessed at: 24.04.2025

[22] Pennefather, P. (2023) *Creative Prototyping with Generative AI: Augmenting Creative Workflows with Generative AI (Design Thinking)*, 1st ed. APress,

[23] Paneru,Pip., Paneru, Bish., Poudyal, R., Shah, K. (2024) *Exploring the Nexus of User Interface (UI) and User Experience (UX) in the Context of Emerging Trends and Customer Experience, Human Computer Interaction, Applications of Artificial Intelligence.* Available at: https://doi.org/10.34010/injiiscom.v5i1.12488 Accessed at: 09.05.2025

[24] Norman, D., Nielsen, J. (1998). *The Definition of User Experience (UX).* Available at https://www.nngroup.com/articles/definition-user-experience/ Accessed at: 09.05.2025

[25] Nielsen, J., Molich, R. (1990). *Heuristic Evaluation of User Interfaces*. Available at: https://dl.acm.org/doi/10.1145/97243.97281 Accessed at: 10.05.2025

[26] The SAGE Handbook of Qualitative Research Design, (2022) 2 volume set, Chapter 24, pp. 399-413

[27] Lim, W., Gunasekara, A., Pallant, J., Pallant, J., Pechenkina, E. (2023). *Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators.* https://doi.org/10.1016/j.ijme.2023.100790 Accessed at: 13.05.2025

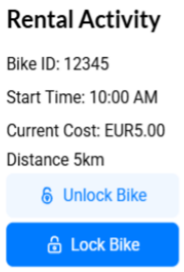
Appendix

I. Participant Evaluation Survey Questions

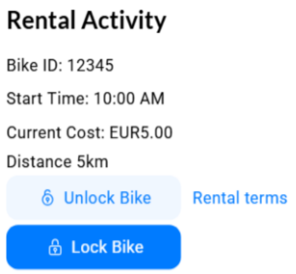
The following questions regarding participants' prior experience with digital tools were asked before the testing session:

1. How often do you use mobile apps for transportation (e.g., ride-sharing, bike-sharing, public transit apps)?
2. On a scale from 1 to 5, how comfortable are you with using new digital apps without instructions?
3. How do you typically learn to use a new app—by exploring it yourself, following tutorials, or asking for help?

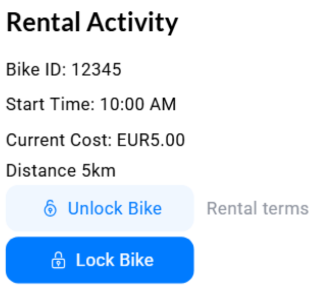
II. Evolution of Improvements on the Rental Activation Screen



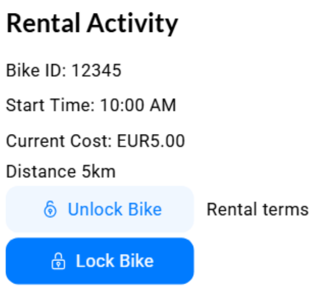
As-Is



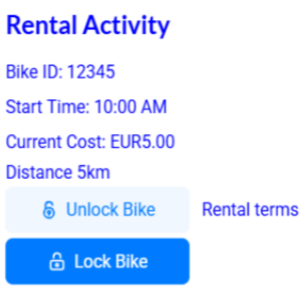
Iteration 1: Text “Rental terms” created but with wrong color.



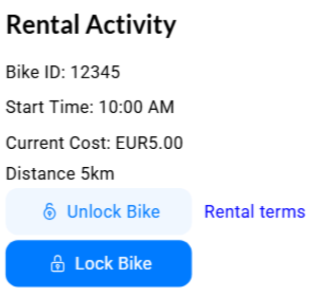
Iteration2: Change color of “Rental terms” text. Result was wrong color.



Iteration 3: Change color of “Rental terms” text. Result again was wrong color.



Iteration 4: Change color of “Rental terms” text.Result was expected but unexpected color change for other text.



Iteration 5: Change color of other text back to initial color. Result as expected.

III. List of Prompts Used in Requirements Implementation

Requirements	Prompts used	Result
1. Enhance the map on the homepage to accurately reflect the geography of Tartu city.	- Create a street map for central Tartu city	Not successful
2. Delete search bar from home page.	- Remove the search bar - Delete the empty space between screen name "Home" and the top of the map	Successful
3. Change the sequence of buttons on rental activation page	- Button with text label Unlock Bike must be positioned above button with text label Lock Bike	Successful
4. Switch the highlight colors between buttons on rental activation page	- Modify the UI design by switching the background colors and font colors of the buttons. Maintain the existing layout. Ensure that the following elements remain unchanged: Button labels: "Unlock Bike" and "Lock Bike." Font size and style. Apply the color swap consistently, ensuring readability and accessibility. Use contrast ratios that meet accessibility guidelines. - For buttons Unlock Bike and Lock Bike - switch the background colors and font colors of the buttons.	Not successful
5. Create a link from rental activation page what navigates to Support page	- Create a link next to Unlock Bike button with text: Rental terms - Change the text color to match with the text color of other text on that page. - Change the text color to black. - Change the text color to blue - Change text color of Rental terms text to blue - For text Rental terms - change font color to the same color as the color of text label on button Unlock bike.	Successful

IV. Participant Evaluation Consent Form

Researcher: Mattias-Silvester Mustonen
Institution: University of Tartu
Purpose: This participant evaluation session is conducted as part of Mattias-Silvester Mustonen's Master's Thesis research at the University of Tartu. The goal is to analyze user interactions to improve usability.

Recording and Data Usage

- The session will be recorded, including **computer screen activity, voice, and facial movements**.
- The recording will be securely stored in **Mattias-Silvester Mustonen's Google Drive** until the thesis is defended, but no later than **December 31, 2025**.
- The recordings will be used **exclusively for educational purposes** and will not be transferred to any third party. However, Mattias-Silvester Mustonen may **show the recording as evidence to third parties** involved in the thesis evaluation.
- During the thesis writing process, any potential identifiers will be either removed or aggregated to ensure confidentiality, and all data will be presented in an anonymized format.
- Participation is **voluntary**, and the participant can withdraw at any time before the analysis begins.

Consent Statement

I, the undersigned, confirm that:

- I have read and understood the purpose of this participant evaluation test.
- I agree with the recording of my screen, voice, and facial movements.
- I understand that my data will only be used for research purposes and stored securely until December 31, 2025.

- I acknowledge that my recordings may be shown as evidence to third parties but will not be transferred to them.
- I understand that I can withdraw from the study at any time before the analysis starts.

Participant's Personal Details:

Date: _______________

Participant's Signature:

Researcher's Signature:

V. List of Activities Performed to Gather Information about AI Applications

The activities undertaken to gather relevant AI supported applications:

1. Search Engine Utilization:

   1.1 Google was employed as the primary search engine to identify tools and resources.

The search was conducted using the following keywords:

   1.1.1 best AI tools for UX

   1.1.2 AI prototyping tools

2. AI-Based Query:

   1.2 ChatGPT was utilized to enhance the search process. The used prompt was:

   List me Integrated AI-based Prototyping Solutions that can create prototypes directly from text. Provide a table.

VI. Prompt for Creating Bike Sharing App Prototype in Visily

To initiate the prototyping process in Visily, the following prompt was employed:

Create a high fidelity fully interactive prototype for bike sharing application. The application interface should include following key features:

User registration and payment setup

- Show registration form where users can enter personal details: name, email, phone number.
- Enable saving credit card details as payment method for future use.

Landing page with city map and bike availability

- Display a city map of Tartu.
- Show bike stations with real-time availability of bikes.
- Provide interactive markers on the map that users can click to get details about each station.

Rental activation and management

- Users can start or end a rental by interacting with the interface by tapping call-to-action buttons to lock or unlock a bike at rental stations.
- Show a section displaying ongoing rental activity, including bike id, start time, distance in kilometers and cost estimation in euros.

Bike rentals and payment history

- A dedicated page or section where users can view their past bike rentals.
- Show payment details of each rental.

Support and help

- A page with the service provider's contact number, email address, and contact form is available.
- A page or section with FAQ about rental terms, payments and usage of the application.

Usability and design

- User interface has a clean minimalistic intuitive design.
- All pages and sections are interactive and linked with hotspots.
- Prototype is following usability heuristics: match between system and the real world, clear navigation between pages and sections, consistency across pages and interactions, visibility of system status, error prevention, user control.

Based on prompt and requirements the flow incorporates 5 screens.

VII. Prototype Creating process in Visily

The Bike Sharing application prototype was developed using the Visily Pro version, utilizing the prompt specified in Appendix VI. The Pro version allows for an input of up to 2,000 characters, compared to the 500-character limitation of the free version. This extended capacity was sufficient to accommodate the 1,737-character prompt used in this experiment. Additionally, the Pro version offers enhanced collaboration features and provides 3,000 AI credits per month, whereas the free version includes only 200 AI credits, which would have been insufficient for comprehensive experimentation with the application's capabilities.

The process of learning and utilizing the AI-driven functionalities of the application was intuitive and efficient. A total of five iterations were required to identify the most effective approach for achieving optimal results. With each iteration, the AI tool demonstrated improved contextualization, refining the design by incorporating relevant images and aligning textual content with the bike-sharing context. The most efficient method for generating the final prototype was found to be constructing the entire prototype from the initial prompt. The result was highly impacted on prompt quality and screens' description quality. During the initial iterations, it was observed that Visily automatically generated a launch screen and included it in the total screen count. To address the issue where the requirement for five pages was condensed into four screens—thereby disrupting the initial specifications—the total number of screens in the screens' description process in Visily was increased to six. While editing the generated design using the AI tool—still in its beta phase—was feasible, each query required approximately 30 to 60 seconds to process, making the editing process time-consuming. However, essential modifications, such as adjusting element alignments and amending textual content, were effectively performed through text-based prompts. Although the addition of icons and other elements via text prompts to existing pages was possible, Visily was unable to consistently maintain the established design style across all screens. Nevertheless, consistency was achievable by replicating one of the AI-generated pages as a design template and subsequently modifying its content.

The AI also generated interactions between hotspots and screens alongside the screen designs. However, not all interactions between elements were established, and the AI did not provide functionality for editing specific interactions through text-based prompts. Instead, the only available option was to overwrite all existing interactions using AI, which still did not yield fully accurate results. Ultimately, achieving the correct interactions between elements and screens required 4 manual adjustments. From a time-efficiency perspective, this limitation did not present a significant challenge, as manual editing was both intuitive and straightforward.

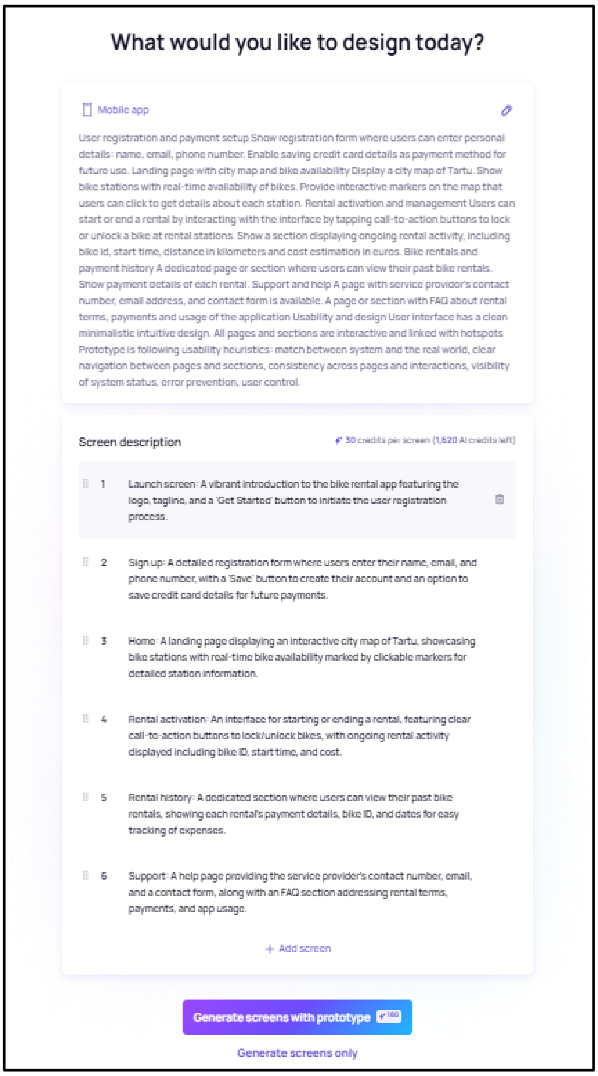


Figure 13 illustrates the AI-generated screen descriptions

The AI-assisted prototyping functionality within Visily was found to be highly accessible and intuitively integrated into the application interface. The process commenced with inputting a predefined prompt, as outlined earlier in this chapter, into Visily's AI-powered input section. A mobile application prototype was selected, based on the requirement that the majority of bike rentals would be conducted via mobile devices. Figure 13 illustrates the AI-generated screen descriptions derived from the initial prompt. At this stage, it is possible to modify both the number of screens and the prompt itself, a step that is highly recommended to prevent time-consuming design modifications and replication efforts in later phases. In this experiment, the number of screens was manually adjusted to 6, as Visily automatically includes a launch screen, which resulted in the exclusion of one required screen with essential functionality. The names of created 6 screens were: Launch Screen, Sigh-up, Home, Rental Activation, Rental History, Support.The subsequent phase involved theme creation. As previously discussed, multiple options were available for this process. In this instance, theme generation from text was utilized, incorporating sections from the initial prompt that outlined usability and design principles. Figure 14 presents an illustration of the theme selection process.

The AI-generated prototype demonstrates alignment with several key usability heuristics, including visibility of system status and overall design consistency. The interface exhibits an intuitive and user-friendly structure. However, certain limitations were identified, particularly in navigation, as the back buttons were non-functional, leading to dead-end situations. Due to the absence of back-end functionality, heuristics related to error prevention and user control could not be comprehensively evaluated.

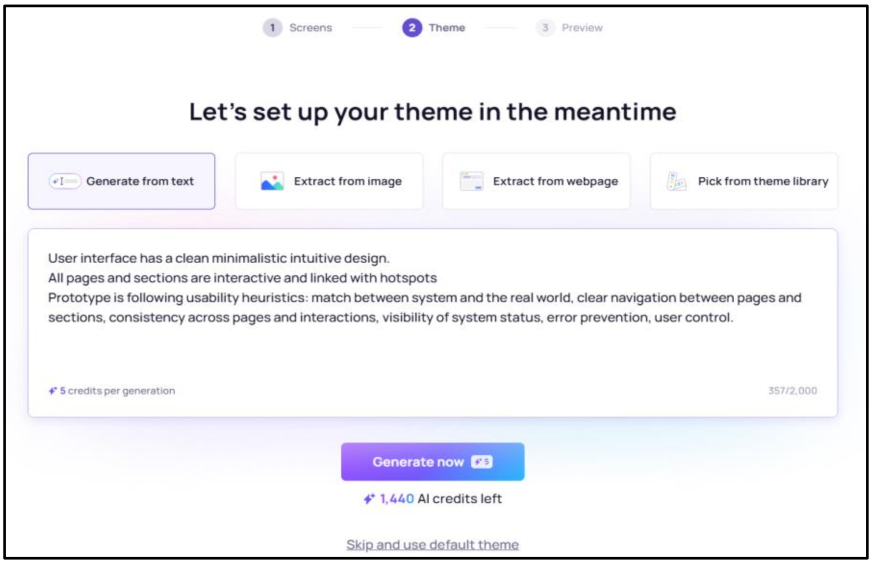


Figure 14. Theme creation in Visily

The heuristic of matching system status with the real world was only partially achieved, as the displayed currency did not correspond to the explicitly stated requirement of using Euros. A notable observation was the automatic generation of small bicycle images, which may indicate the self-learning capabilities of Visily's AI. This feature was not present in earlier iterations but appeared after an iteration in which the theme was extracted from Nextbike[10], a commercial bike rental company. This iteration was part of the learning process of the application, where different functionalities of Visily were experimented. The theme extraction is Visily's functionality which enables users to enter design references. Figure 15 presents the initial prototype.

[10] https://www.nextbike.de/en/

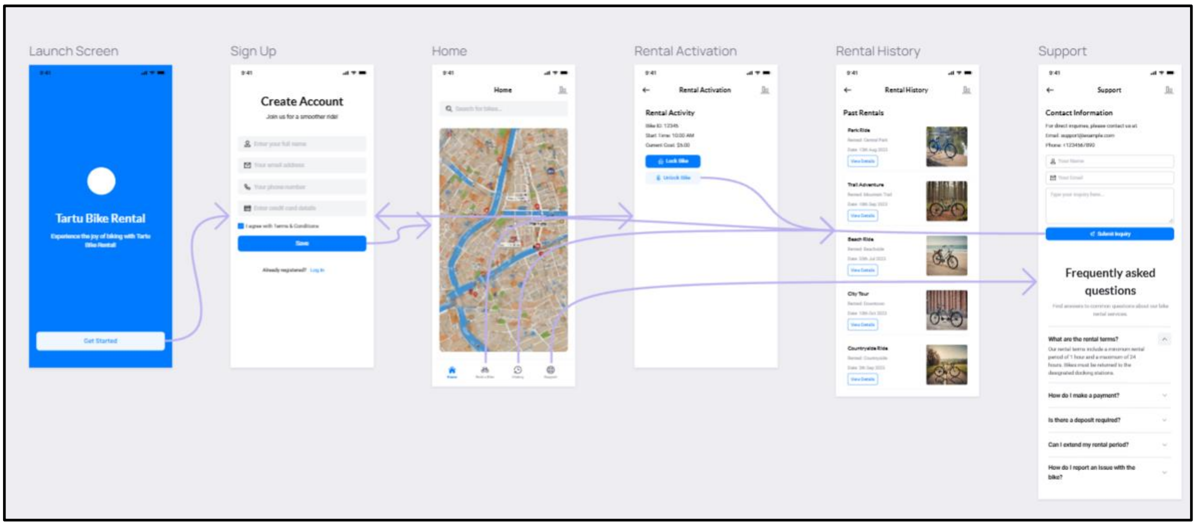


Figure 15. Initial prototype in Visily

To improve navigation between pages, three iterations with the same prompt were conducted using an AI feature Generate prototype to re-generate interactions. While these iterations resulted in noticeable improvements, some minor interactions, such as connecting back buttons with previous pages, required manual adjustments to ensure the overall usability of the prototype.

In the subsequent step, textual modifications were implemented, such as updating button labels on the Rental History screen and adjusting the currency display on the Rental Activation page. These modifications were executed seamlessly within a single iteration. However, inserting a missing text line on the Rental Activation screen required three iterations, as it initially affected the alignment of the buttons. By refining the prompts to specifically address the alignment issue, Visily successfully resolved the problem. Figure 16 presents an illustration of the alignment prompt and the corresponding result.

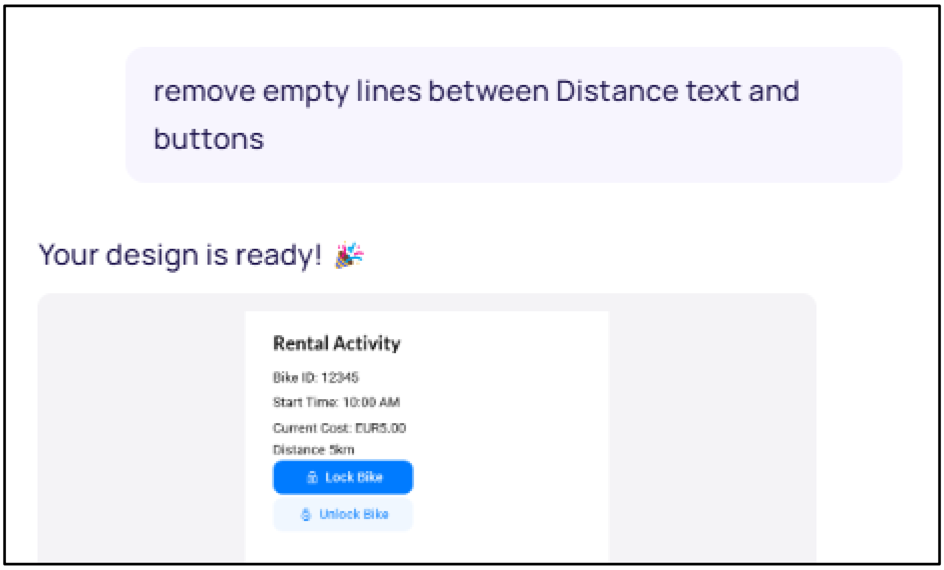


Figure 16. Illustration of a Prompt and the Result Generated by Visily's AI-Powered "Modify Existing Design" Feature

The finalized prototype, intended for user testing, is accessible on Visily's collaboration platform[11].

[11] Link to Visily prototype

VIII. Participant evaluation follow-up questions

1. How would you describe your overall experience with the prototype?
2. What, if anything, frustrated or confused you while completing the tasks?
3. Were there any parts of the prototype that felt particularly easy or intuitive to use? Why?
4. Did the prototype work as you expected? If not, what would you change?
5. If you could improve one thing about this prototype, what would it be?

Licence

Non-exclusive licence to reproduce the thesis and make the thesis public

I, Mattias-Silvester Mustonen ,

(author´s name)

grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the digital archives of the University of Tartu until the expiry of the term of copyright, my thesis

AI-enabled prototyping tools in user-centered design ,

supervised by Grace Eden, Yana Halas ;

(supervisor´s name)

2. grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright;

3. am aware of the fact that the author retains the rights specified in points 1 and 2;

4. confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Mattias-Silvester Mustonen

14/05/2025