An Approach to Assist the Design of User-Centered Solutions based on Generative AI

Thiago Queiroz
Institute of Computing
Federal University of Amazonas
Manaus, Amazonas, Brazil
thiago.queiroz@icomp.ufam.edu.br

Gretchen T. de Macedo
Institute of Computing
Federal University of Amazonas
Manaus, Amazonas, Brazil
gtmacedo@icomp.ufam.edu.br

Igor Steinmacher
School of Informatics, Computing and Cyber Systems
Northern Arizona University
Flagstaff, Arizona, USA
igor.steinmacher@nau.edu

Tayana Conte
Institute of Computing
Federal University of Amazonas
Manaus, Amazonas, Brazil
tayana@icomp.ufam.edu.br

Ana Carolina Oran
Institute of Computing
Federal University of Amazonas
Manaus, Amazonas, Brazil
ana.oran@icomp.ufam.edu.br

Bruno Gadelha
Institute of Computing
Federal University of Amazonas
Manaus, Amazonas, Brazil
bruno@icomp.ufam.edu.br

ABSTRACT

User-Centered Design (UCD) aims to ensure software quality by aligning its development with user needs, enabling users to guide project decisions and contribute to the overall quality of the product. Some methodologies, such as Design Thinking and Lean UX, incorporate UCD, but they often lack a rigorous structure that focuses on the rapid validation of solutions. This paper investigates the Design Sprint, a UCD approach, which offers short development cycles for validating solutions with the users. We developed an approach based on generative Artificial Intelligence (AI) to automate Design Sprint activities through structured prompts designed with prompt engineering techniques. We evaluated the proposed approach through a feasibility study involving 38 participants. The study included introductory training and the practical application of the prompts in team-defined projects. Results indicate a predominantly positive reception of the proposed approach, with 85,7% of the teams rating all artifact-generating activities as either "very helpful" or "helpful" during the Design Sprint phases of understanding, sketching, decision-making, and prototyping, which reinforces the feasibility of the approach presented in this study.

KEYWORDS

User-centered Design, Generative AI, Design Sprint, Agile methodology

1 Introduction

User-Centered Design (UCD) is a software development approach that prioritizes users' needs by placing them at the center of the creation process. Its goal is to understand the contexts of use, objectives, and key challenges faced by users, enabling solutions that are better aligned with their expectations [1]. According to ISO 9241-210:2019, it is an iterative process that emphasizes users' needs through continuous cycles of validation and refinement based on the solution under development [8].

Several methodologies implement UCD. One of them is Design Thinking, an iterative, non-linear, and innovation-oriented process structured into five phases: empathize, define, ideate, prototype, and test [14]. Its application is aimed at contexts of uncertainty and poorly defined problems, in which the exploration of creative solutions is required. Another approach is Lean UX, which derives from the principles of Lean Startup and emphasizes agility in delivering value to users through short validation cycles, minimal documentation, and strong collaboration among product, design, and development teams [6].

In contrast, Design Sprint [11] incorporates elements from these approaches but differs by presenting a systematized structure oriented toward the rapid validation of solutions. Developed by Google Ventures[1], this process is conducted over five days, with stages that include understand, diverge (sketch), converge (decide), prototype, and test.

Recent advances in generative Artificial Intelligence (AI) have driven the adoption of the Large Language Model (LLM) as a tool to automate tasks in software engineering. These models demonstrate the ability to interpret diverse contexts and generate contextually relevant responses, making them valuable in various creative processes. This potential has motivated a series of studies investigating their integration into software engineering practices [5].

Recent papers have already explored the use of AI in the application of the Design Sprint. Notable examples include the collaborative creation of board games [18] and the development of a bot for the Slack platform[2], designed to facilitate the execution of the methodology [15]. However, both studies present significant limitations: the first is restricted to a specific context and employs a customized version of the Design Sprint, while the second, aimed at generating pitch materials, relies on the Slack platform and also uses an adapted version of the Design Sprint. On the other hand, the present work proposes the use of AI to provide more generalizable support for the Design Sprint, while preserving its methodological structure.

The application of the Design Sprint presents challenges that may limit its adoption in specific organizational contexts. Key obstacles include: (i) dependence on the presence of industry experts, whose participation is crucial for knowledge sharing but is not always feasible; (ii) the need to produce multiple artifacts within a constrained time frame, such as solution map, solution sketches,

[1]https://www.gv.com/
[2]https://slack.com/

and prototypes, which requires time and design skills and may limit teams with fewer available resources; and (iii) the requirement of a facilitator, whose responsibility is to ensure the process dynamics, mediate discussions, and guarantee the completion of tasks, its absence or inadequate execution tends to impact activities.

In this context, there is an opportunity to explore the potential of AI, particularly LLMs, as a resource to mitigate these challenges. These models can (i) fill knowledge gaps when expert presence is not possible, (ii) support the agile production of artifacts, and (iii) even act as facilitators within Design Sprint activities.

Given this characterization of the problem and the identified gap in the literature, the following research question is formulated to guide this investigation: How can LLMs support the phases of the Design Sprint?

To address this research question, a feasibility study was conducted with 38 computing students, organized into seven teams. Each team employed a set of prompts developed through prompt engineering based on the revised 2024 version of the Design Sprint. This version, made available on the Miro platform[3] by Jake Knapp [10], incorporates updated techniques and instructions. The study leveraged ChatGPT[4] to support activities ranging from ideation to prototyping the solution. As a contribution, this work provides evidence of the feasibility of applying an AI-assisted Design Sprint in the phases of understanding, diverging (sketching), converging (deciding), and prototyping.

2 Background

This section presents the theoretical foundations of UCD and Design Sprint, intending to contextualize the main topics addressed in this research and discuss related work.

2.1 User-Centered Design (UCD)

UCD is a design approach that aims to develop solutions aligned with users' needs, with active user participation as its fundamental principle [1]. In this context, users are continuously involved, providing feedback that guides the development of the solution [8].

Among the main advantages of this approach, the following stand out: (i) increased productivity, as continuous user involvement enables the early identification of discrepancies, preventing errors from advancing to later project stages and causing rework; (ii) greater alignment of the solution with users' real problems, since they act as a "compass" that guides the project in the right direction; and (iii) enhanced product success, achieved through improved usability and user satisfaction resulting from continuous evaluation and iterative refinement based on identified needs [1].

Despite its benefits, the primary challenge of UCD lies in its collaborative nature. The need for frequent interaction with multiple stakeholders may lead to additional costs, such as financial, temporal, and logistical costs, as well as impose demands on human resources that are not always available when needed. Moreover, it is necessary to properly manage the level of user involvement to balance development speed with its participation [7].

Several methodologies have been developed based on UCD principles, illustrating both the strengths and challenges of this approach. Design Thinking is one such approach, recognized for its flexibility and innovation orientation [14]. It is considered effective in contexts of uncertainty and poorly defined problems, as it encourages co-creation and continuous experimentation. However, it often lacks strict time constraints for its activities.

Lean UX, on the other hand, prioritizes rapid delivery and continuous validation with users through short feedback cycles, promoting strong collaboration among stakeholders [6]. Nevertheless, the frequency of interactions with users over time requires constant availability and engagement, which may be unfeasible when stakeholder access is limited.

Due to these limitations, which stem from the lack of a rigorous schedule in the mentioned methodologies, Design Sprint emerges as a promising alternative. It is an intensive and highly organized methodology structured within a five-day schedule that enables the development and validation of solutions in short cycles. This structure allows the combination of an agile development pace with user involvement throughout the process.

2.2 Design Sprint

Although several methodologies adopt UCD principles, many of them involve long iteration cycles and extensive schedules. In this context, the Design Sprint stands out for its structured approach to rapid solution validation, enabling teams to test ideas with users within a short time frame.

The Design Sprint integrates elements of Design Thinking, agile methodologies [19], and Lean UX [6], organizing the process into five days accompanied by detailed checklists with a strict schedule and predefined tasks that guide activities within specific time frames. This structure highlights the highly organized nature of the methodology, which is capable of guiding teams throughout the project. The five phases of the Design Sprint are described below:

(1) Understand - The first day is dedicated to defining the problem to be addressed. The long-term goals, metrics, and risks are established. The team constructs a map representing users' interactions with the product and, with the support of industry experts, refines the initial assumptions. Then, the focus of the sprint is chosen, for example, the core functionality to be tested. By the end of the day, all team members are aligned regarding the ideas, including the sprint's focus [22].
(2) Diverge (Sketch) - On the second day, the objective is to produce sketches to identify potential solutions. The team seeks inspiration from existing ideas and performs a "remix" of them to propose innovative alternatives. Subsequently, sketches are created to transform abstract concepts into concrete representations. At the end of the day, all sketches are collected and organized for the decision-making phase. Additionally, the process of recruiting users who will participate in the testing phase begins [24].
(3) Converge (Decide) - The third day focuses on selecting the best proposals to be included in the prototype. The sketched ideas are analyzed, discussed, and criticized, and the most promising solutions are selected. The team then builds a

[3]https://miro.com/miroverse/the-2024-design-sprint/
[4]https://chatgpt.com/

storyboard filled with the selected sketches, which serves as the foundation for prototype development [25].

(4) Prototype - On the fourth day, the goal is to build a prototype. To achieve this, tools that prioritize speed over visual quality are employed. The workload is distributed among team members, and after the prototype is completed, it is tested and adjusted internally. By the end of the day, the prototype is ready to be tested by users [23].

(5) Test - On the fifth and final day, the prototype is evaluated by users. Interviews are conducted by a moderator while the rest of the team observes the sessions remotely, taking real-time notes. After the interviews, observations are gathered and analyzed to identify patterns and determine whether the sprint's objectives have been achieved. Based on this analysis, the team decides whether to conclude the sprint or initiate a new one [21].

The outcomes obtained at the end of a sprint can be viewed in two ways: as a "failed success," in which the prototype is well received but requires improvements, or as an "efficient failure," where the solution proves inadequate but did not demand high resource investment to reach that conclusion. Regardless of the result, the team obtains a tested prototype that generates valuable data to guide the next steps of the project. Therefore, the Design Sprint enables teams to develop and validate solutions rapidly, reducing risks and shortening development cycles.

2.3 Related Work

Considering the growing number of initiatives in the field of software engineering aimed at applying generative AI to activities related to user experience [12], recent studies have investigated the use of this technology in practices associated with UCD.

Among these initiatives, Torii et al. [18] proposed the collaborative creation of board games through the application of an AI-assisted Design Sprint. However, as the study was conducted within a specific context, it presents limitations regarding the generalizability of its approach to other software engineering scenarios, as well as the use of a customized version of the Design Sprint. In contrast, the present research seeks to develop a generalizable approach.

Another relevant work is that of Rana and Cheok [15], who proposed the development of a bot for the Slack platform designed to act as a facilitator for a five-stage Design Sprint. Despite its similarity to the present study in focusing on process automation, it is worth noting that the original methodology was modified by the authors and reorganized into five new stages: ethnography, divergent ideation, convergent distillation, strategic visualization, and pitch material production for the validation phase. The distinguishing feature of the present research lies in preserving the original methodological structure of the Design Sprint.

Ding et al. [3] propose DesignGPT, a framework grounded in Design Thinking principles and the design process, in which multiple AI agents simulate different roles within a team to support designers. Although it shares with the present work the goal of employing AI to assist the design process, the methodological approaches diverge. While DesignGPT explores multi-agent collaboration in Design Thinking, this research is based on the use of the Design Sprint without involving multiple agents.

Chen et al. [2] investigate how AI can support humans in conceptual design, considering stages such as problem definition, idea generation, idea selection, idea evaluation, and idea evolution. The authors highlight that AI proved especially helpful in the stages of problem definition and idea generation, and to a lesser extent in idea evolution, while idea selection and idea evaluation remained primarily human responsibilities. Nevertheless, the presence of AI led participants to conduct these stages more carefully, as they needed to engage critically with its generated responses. These findings align with the purpose of the present research, emphasizing the collaborative role of AI in the creative design process and the integration between human critical intervention and AI-driven creativity. However, the studies differ in their approach to interactions: while Chen et al. employ free-form prompting, the present work proposes a structured sequence of prompts, aiming for greater methodological rigor and replicability, while incorporating the Design Sprint framework.

Ekvall and Winnberg [4] propose promptBoard, an interactive interface that integrates ChatGPT into the User Experience (UX) design process. The tool supports the ideation and prototyping phases of chatbot design by automating the sending of structured prompts to the model's Application Programming Interface (API) through a web application. Although its scope is limited to chatbot design, the study aligns with the present research in exploring AI as a support tool for design activities. Among the differences, the mode of interaction stands out the most. While promptBoard automates the application of prompts, this work performs the process manually, suggesting an opportunity for future research to encapsulate the approach within software.

Although conducted within a Design Sprint workshop, the study by Ekvall and Winnberg does not encompass all Design Sprint activities. The ideation stages, for example, were limited to persona creation, brainstorming, and the How Might We technique. In contrast, the present research systematically addresses the activities of each Design Sprint phase, from ideation to prototyping, becoming more comprehensive and aligned with the original methodology.

The comparative study by Kamnerddee et al. [9] explores AI-assisted Design Thinking, intending to automatically generate a prototype and compare it to one developed by humans. Although the present work focuses on the Design Sprint, both share the objective of supporting a user-centered methodology by automating its stages to enable prototype generation with AI.

In summary, while the related works explore the use of AI in user-centered methodologies, each presents limitations that distinguish them from the present research, such as contextual restrictions involving board games (Torii et al.) or chatbots (Ekvall and Winnberg), customized versions of the Design Sprint (Torii et al.; Rana and Cheok), the use of multiple AI agents (Ding et al.), free-form prompting (Chen et al.), or a focus on Design Thinking (Ding et al.; Kamnerddee et al.). Thus, a gap remains for a systematic and replicable approach that supports teams in executing the Design Sprint with the assistance of generative AI, while preserving its original formulation and applying it to diverse contexts. For this reason, the present study aims to address this gap. It employs the updated 2024 version of the Design Sprint [26], made available by its creator, Jake Knapp, through a template on the Miro platform [5].

3 GenDS

The Generative Design Sprint (GenDS), the artifact developed in this research, is presented as follows: a set of structured prompts designed to adapt the Design Sprint methodology for application with the support of a generative AI, specifically ChatGPT. The artifact aims to automate the execution of the methodology's activities, from the ideation phase to prototyping, while preserving its original structure.

Each prompt contains additional guidelines, including the activity's objective and usage instructions, to help teams use the material effectively. Thus, the proposal seeks to support teams in executing the AI-assisted methodology by providing precise and structured instructions for each sprint activity.

The development of the prompts was guided only by the first four phases of the Design Sprint, as described in the book "Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days" [11]. For this purpose, an exclusive template of the methodology, available on the Miro platform and developed by Jake Knapp, the creator of the Design Sprint [11], was employed. This material, updated in 2024, incorporates revised techniques compared to the original version of the book published in 2016. For each phase, specific prompts were developed based on prompt engineering techniques, providing the AI with clear instructions to generate responses aligned with the goals of each activity. Each prompt also includes supplementary explanations, such as a description of the activity's purpose, usage instructions, and guidelines for evaluating the AI-generated output.

Additionally, a **Pre-Sprint** phase was introduced before the main sprint, aiming to prepare the model for the subsequent activities. In this stage, ChatGPT is informed that the conversation context is about Design Sprint focused on prototype creation and is instructed to adopt a specific role within the process. Furthermore, since defining a big challenge is a key prerequisite for initiating the sprint [11], a prompt was developed to formulate this initial challenge, serving as the starting point for the structured sequence of methodology phases.

During development, each prompt was organized sequentially, following the order of activities provided by the methodology, as the activities are strongly dependent on one another. The following guidelines were adopted with a focus on the practical applicability of the prompts and clear communication for participants:

(i) division by phase;
(ii) activity's number;
(iii) activity's title;
(iv) objective;
(v) pre-prompt instructions;
(vi) the prompt to be submitted to ChatGPT;
(vii) post-prompt instructions;
(viii) description of the expected response;
(ix) manual activities, when applicable; and
(x) evaluation questionnaire, when necessary.

[5]https://miro.com

Regarding item (vi), the prompt to be submitted to ChatGPT maintained a standardized and consistent structure across all activities. This standardization was intended to facilitate participants' familiarity with the prompt format and to guide the model's ability to adapt to the instructions provided at each stage of the methodology.

Figure 1 partially presents a segment of activity four, illustrating the structure of the material. The complete set of prompts used in this study is available as supplementary material[6].

Day 1: Understand the Problem

4. **Activity 4 (Define the Long-Term Goal)** - 🚩 **Includes an evaluation questionnaire at the end**

- **Objective:** Present the big challenge chosen in Activity 3 to ChatGPT so that it can generate suggestions for long-term goals for the sprint.
- **Pre-prompt instructions:**
  1. **Attention:** This activity includes a prompt that must be customized by your team! Read the second instruction below carefully for more information.
  2. Modify and insert the prompt highlighted in yellow into ChatGPT (see ➢ **Prompt to be submitted to ChatGPT**), filling in the section highlighted in red with your system's context.
  3. Then, wait for ChatGPT's response before proceeding to the next activity.

➢ **Prompt to be submitted to ChatGPT:**

[Title 1]
Day 1: Understand the Problem

[Título 2]
Suggestions for Long-Term Goals for the Sprint

[Instruction]
Based on the challenge provided in the content below, adopt a bold, optimistic, and even idealistic approach.
Rules for the goals:
- Write one sentence describing a perfect outcome for customers.
- Do not use numbers or metrics
- Do not prescribe a solution.

Then, generate a numbered list with as many broad and inspiring long-term goal sentences for the product as possible.

[Reflection]
To guide your formulation, reflect on:
- Why are we doing this project?
- Where do we want to be in 6 months, 1 year, or even 5 years?

[Context]
⚠️ When filling in, remove the brackets and this note.
[Insert here the big challenge chosen in **ACTIVITY 3**]

[Expected Output]
- Display Titles 1 and 2 prominently, using a large font size.
- Generate a numbered list of suggested long-term goals.
- The goals should be based on the provided content and formulated in a clear, ambitious, and inspiring manner.

Figure 1: Partial example of an activity.

The prompts were organized into activities distributed across the methodology's phases, as follows: three activities in the **Pre-Sprint** phase, twelve in the **Understand** phase, two in the **Diverge (Sketch)** phase, three in the **Converge (Decide)** phase, and one in the **Prototype** phase. Each of these activities is listed in Table 1, along with information regarding their respective objectives and generated artifacts. The proposed approach does not include the user testing phase, since the research focus lies in automating the steps up to prototyping.

[6]See Section *ARTIFACT AVAILABILITY*

The division of activities by phase aims to facilitate the use of the material by offering guidance in a structured and sequential manner. It is noteworthy that the **Understand** phase contains the most significant number of activities, which can be explained by the high volume of tasks related to ideation and brainstorming, both key steps of the first day of the Design Sprint methodology.

Figure 2 presents the synthesized Business Process Model and Notation (BPMN) flowchart of the proposed approach, highlighting the input data and the artifacts generated in each phase of the activities described in Table 1.

Adaptations were necessary to enable the automation of activities that are traditionally collaborative and manually conducted. In the standard format of the methodology, most activities rely on the use of physical Post-its. For this research, these dynamics were adapted into a textual workflow through the use of prompts, designed to preserve the core principles of the Design Sprint.

To facilitate the use of the artifact, all the prompts were grouped into a single document, organized into sections identified by phase and activity. Thus, the structured set of prompts represents an initial proposal for automating the Design Sprint with the support of generative AI, assigning the model the role of facilitator throughout the process.

4 Feasibility Study

To evaluate the proposed artifact, a feasibility study was conducted focusing on the practical application of the prompts throughout the phases of the Design Sprint, with the purpose of enabling a critical analysis of the impact of automation on each of its activities.

4.1 Study Context

The study was conducted over a 14-day period within the context of the Software Engineering Practice course at the Federal University of Amazonas, where students were required to develop a system of their choice from scratch. A total of 38 students from the Institute of Computing participated, organized into seven teams: four teams of six members, two teams of five members, and one team of four members. It is noteworthy that all teams used the free version of ChatGPT during the activities. All participants signed an informed consent form authorizing the collection and analysis of their responses.

4.2 Study Steps

The study was structured into three phases: (i) an introductory training session on the Design Sprint in its standard format, lasting approximately two hours; (ii) a second training session focusing on the use of generative AI as support for the methodology, also lasting about two hours; and (iii) the practical application of the prompts in the context of a system chosen by each team.

In the first phase, a classroom training session was conducted to introduce participants to the fundamental principles of the Design Sprint, ensuring a shared conceptual understanding among them. In the second phase, a practical simulation was presented, also in class, demonstrating the application of the proposed prompts in a hypothetical scenario involving the development of a travel application similar to Uber. The exercise consisted of applying the methodology in a guided manner, starting from an initial concept and progressing through the generation of necessary artifacts until a prototype of the solution was obtained. The presentations used during the training sessions are available as supplementary material[7].

Finally, in the third phase, the teams were instructed to apply the prompts to the system they had selected, with the goal of generating insights. It is worth noting that, within the course context, the students were preparing to begin the first sprint of the Scrum framework. Thus, the implementation of the Design Sprint was conceived as an initial strategic activity aimed at identifying key aspects of the system prior to the start of the Scrum.

4.3 Practical Application of GenDS

Each team used the same document, which contained a preliminary set of structured prompts designed to automate the stages of the Design Sprint. Based on this material, the teams carried out a series of activities with the support of ChatGPT, encompassing the process from the ideation phase to the generation of the proposed solution prototype.

Each activity included a specific prompt containing its objective and highlighted sections that each team was required to edit with information related to their chosen system. Furthermore, at the end of each activity, instructions were provided regarding the generated response, including a description of an "expected ChatGPT response," which aimed to help teams verify whether the prompt functioned as intended. The examples used during the classroom training sessions were also made available as additional references for this verification process.

Teams were instructed to review each response generated by ChatGPT manually, thereby promoting a collaborative interaction between human intervention and the model's creativity to ensure contextual adequacy. This process allowed responses to be used fully or partially, rewritten, regenerated, manually completed, or, if necessary, entirely replaced. Similarly, the provided prompts could be modified by the teams according to their needs, offering flexibility in their use, enabling the documentation of improvement suggestions for the research, and encouraging participants' critical engagement with each activity.

4.4 Organization of Activities by GenDS Stages

The flow of the first four stages of the GenDS as applied in the study:

- **Day 1 - Understand:** In this phase, the teams used ChatGPT to create the initial Design Sprint artifacts. Based on the big challenge selected according to each team's chosen system, participants requested the generation of suggestions for long-term goals. Subsequently, with the support of ChatGPT, suggestions for metrics related to these goals and associated risks were generated. The model was also used to produce How Might We (HMW) notes, which aimed to identify problems and stimulate reflections on possible solutions. These notes were later organized into a solution map at the end of the phase. Additionally, ChatGPT generated suggestions for user profiles, their initial steps, the expected core experience, and the expected outcome. The phase concluded with the

[7]See Section *ARTIFACT AVAILABILITY*

Table 1: Activities automated by the proposed prompts, with their respective objectives and generated artifacts.

Day	Activity	Objective	Generated Artifact
Pre-Sprint	A1	Introduce the Dynamic to ChatGPT	Not apply*
	A2	Give a Persona to ChatGPT	Not apply*
	A3	Choose a big Challenge	Big challenges
Understand	A4	Set a Long-Term Goal	Long-term goals
	A5	Define Metrics	Metrics
	A6	Define Risks	Risks
	A7	Summarize the Chosen Long-Term Goal, Metrics, and Risks	Not apply*
	A8	Define HMW Notes	HMW notes
	A9	Organize HMW Notes	Organized HMW notes
	A10	List Customers	Customer types
	A11	Define the Customers' Initial Steps	Customers' initial steps
	A12	Define the Customers' Core Experience	Customers' main steps
	A13	Summarize the Chosen Customers and Interaction Flows with the Product	Not apply*
	A14	Define the Expected Outcome for the Solution Map	Expected outcomes
	A15	Make a Solution Map	Product interaction map
Diverge (Sketch)	A16	Crazy 8s	Eight solution variations
	A17	Solution Sketch	Detailed solution
Converge (Decide)	A18	Summarize the Solutions to be Prototyped and Tested	Not apply*
	A19	Storyboard Flow	Prototype's first six steps
	A20	Storyboard	Prototype's interactions
Prototype	A21	Prototype	Prototype's screens

*The activity does not involve the generation of directly measurable artifacts but rather produces intermediate commands throughout the process.

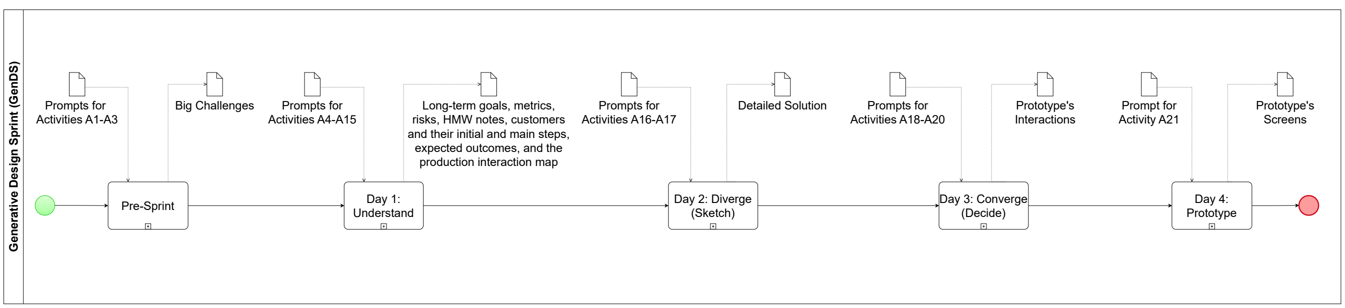

Figure 2: The input data and the generated artifacts of each phase.

creation of the solution map, in which the HMW notes were manually organized into key interactions, assisting in the selection of the sprint's focus.

- **Day 2 - Diverge (Sketch):** The objective of this phase was to explore possible solutions. Teams used ChatGPT to perform the Crazy Eight technique, generating eight variations of initial solutions based on the focus defined in the previous phase. After manually selecting the most promising idea, the model was again employed to generate a three-panel storyboard representing a logical sequence of user interactions with the solution.
- **Day 3 - Converge (Decide):** Based on the previously generated ideas, teams used ChatGPT to expand and detail the interaction flows from the storyboards selected in the previous phase. The model was subsequently used to assist in generating the definitive storyboard for the prototype, guiding the creation of screens in the next phase.
- **Day 4 - Prototype:** In this phase, ChatGPT was employed to generate screen images based on the previously defined storyboard.

4.5 Type of Data Collection

To enable the analysis of the results, each team was requested to share three materials: (i) the complete interaction history with ChatGPT, including the chat content and generated images, both available as supplementary material[8]; (ii) the document containing the structured prompts, which incorporated, in addition to the prompts, questionnaires corresponding to Section 3 of a form, in

[8]See Section *ARTIFACT AVAILABILITY*

which the team evaluated each generated artifact by consensus; and (iii) a completed form aimed at assessing the prompts used.

The decision to include the questionnaires within the prompts document was intended to allow teams to respond to the questions during the execution of each activity, facilitating the recording of real-time perceptions rather than only at the end of the practice.

The form was divided into four sections: (i) the Consent Form; (ii) Participant Profile, aimed at characterizing individual participants; (iii) Team Feedback on the Prompts Used, focusing on the collective perception of the activities done; and (iv) Individual Feedback on the Prompts Used, designed to capture each participant's personal experience during the study.

Sections 3 and 4 of the form primarily consisted of structured questions based on a Likert scale, intended to generate charts that allow evaluation of participants' levels of agreement regarding the different activities. Furthermore, Section 4 included an open-ended question to enable a more in-depth analysis of participants' overall experience with the use of the prompts.

5 Results

This section presents the main results obtained from GenDS in the context of the Design Sprint supported by generative AI.

The data analysis employed a mixed approach, combining quantitative analysis through statistical charts with qualitative analysis based on responses to an open-ended question. This strategy aimed to provide a more comprehensive and in-depth understanding of the participants' opinions about the use of the prompts. The results are presented in the following subsections, accompanied by their respective analyses at both collective and individual levels.

5.1 Collective Analysis by Team

The first level corresponds to the collective analysis conducted by each team. In this case, each group was requested to complete the form collaboratively, expressing a shared perception of the prompts used to generate the artifacts. This approach was designed to capture a collaborative evaluation of each artifact generated based on the prompts, thereby enabling the assessment of each prompt's performance within the context of its corresponding activity.

Figure 3 presents the distribution of the teams' collective responses, considering only the activities for which the prompts generated an artifact. It explains, for instance, the absence of responses related to activities A1 and A2. It shows that, across all 16 artifact-generating activities, the teams' responses were mostly positive, emphasizing the options "very helpful" and "helpful." These results indicate a favorable reception of the study's proposal, suggesting that the prompts supported the teams throughout the AI-assisted Design Sprint.

Additionally, Figure 4 presents the collective responses by team and by activity, allowing for a comparative analysis of prompt performance across different groups. It can be observed that teams T4 and T6 rated all activities as "very helpful," while teams T3 and T7 consistently selected "helpful." Teams T1, T2, and T5, although also reporting predominantly positive evaluations, exhibited some variability in their responses. Notably, only team T5 selected the "neutral" option for any of the activities throughout the process. Overall, the results indicate a predominantly positive reception of the proposed approach, with 85,7% (6 of 7) of the teams rating all artifact-generating activities as either "very helpful" or "helpful," thereby suggesting that GenDS was well received by the teams and perceived as a useful instrument for carrying out the Design Sprint activities supported by generative AI.

5.2 Individual Analysis by Participant

The second level of analysis involved collecting the individual perceptions of each participant. This strategy was conceived to complement the collective analysis, providing greater robustness to the evaluation of the feasibility of using the prompts. Therefore, each participant was asked to complete individual questions within the same form in order to gain a deeper understanding of their personal experience with the prompts used throughout the activities.

Figure 5 reveals that all five questions received widely positive responses, primarily concentrated in the options "strongly agree" and "agree." These results suggest an overall positive experience with the prompts, indicating that they were perceived as useful throughout the process. However, some responses, such as "neither agree nor disagree," appeared in questions Q2, Q3, Q4, and Q5, as well as "disagree" in Q4. These responses suggest that some participants experienced certain limitations or difficulties during the activity.

Figure 6 shows that participants' expectations regarding the use of prompts were, in general, met in a predominantly positive manner.

The answers to the open-ended question revealed different perceptions among participants. Among these, **praise for the GenDS** stood out, especially regarding the clarity provided by the prompts and the additional instructions. There was also **praise for the ChatGPT's performance**, as well as **favorable remarks about the prompts**. On the other hand, some **criticisms of the GenDS** emerged, revealing the need for adjustments to the prompts to improve their suitability for the intended context. In addition, some participants pointed out the need to **enhance the provided material** to make it more user-friendly. Finally, there were also comments concerning the **limitations of image generation** by the model. Examples of each category are presented below.

- Praise for the GenDS:
  - *"It was smooth, the prompts were very didactic regarding their use."* (P23)
  - *"The clear division by days and steps also helped maintain the development pace and avoid creative blocks."* (P28)
- Praise for the ChatGPT's performance:
  - *"It greatly facilitated the process, it is a very practical method that saved the team an amount of time."* (P30)
  - *"[...] greatly increases the speed of the process."* (P2)
- Favorable remarks about the prompts:
  - *"I found the way the chat (ChatGPT) guided its responses according to the requested structures (prompts) very interesting."* (P11)
  - *"Excellent; I was surprised by the capacity of GPT (ChatGPT) when used with the prompts we applied, which further enhanced the tool's capabilities. I will definitely use it even more going forward."* (P3)
- Criticisms of the GenDS:

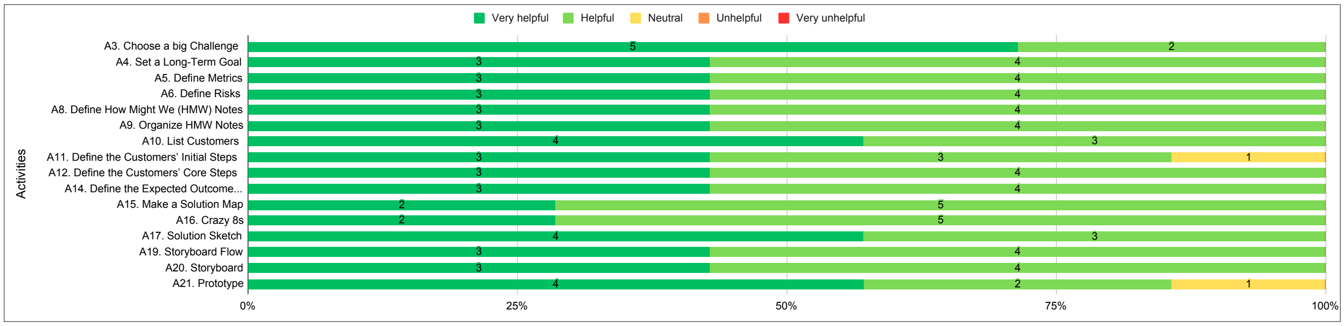

Figure 3: Collective Feedback by Team.

Teams	Activities														
T1	A3	A4	A5	A6	A8	A9	A10	A11	A12	A14	A15	A16	A17	A19	A20	A21
T2	A3	A4	A5	A6	A8	A9	A10	A11	A12	A14	A15	A16	A17	A19	A20	A21
T3	A3	A4	A5	A6	A8	A9	A10	A11	A12	A14	A15	A16	A17	A19	A20	A21
T4	A3	A4	A5	A6	A8	A9	A10	A11	A12	A14	A15	A16	A17	A19	A20	A21
T5	A3	A4	A5	A6	A8	A9	A10	A11	A12	A14	A15	A16	A17	A19	A20	A21
T6	A3	A4	A5	A6	A8	A9	A10	A11	A12	A14	A15	A16	A17	A19	A20	A21
T7	A3	A4	A5	A6	A8	A9	A10	A11	A12	A14	A15	A16	A17	A19	A20	A21

Figure 4: Collective Feedback by Activity.

  - *"[...] some generated responses were generic or disconnected from the proposed context, requiring reformulations and multiple attempts to achieve the desired result."* (P36)
  - *"[...] Generated inappropriate or wrong information, not being precise."* (P4)

- Comments about the material provided:
  - *"[...] I missed more visual examples for some steps."* (P28)
  - *"[...] it (GenDS) was boring because there was too much stuff!"* (P29)
- Limitations of the model:
  - *"[...] Image generation remains problematic when using ChatGPT; sometimes it makes errors or produces incomplete images, requiring requests to be made one at a time."* (P14)
  - *"[...] regarding image generation, even when requesting again with context already established, the expected result was not achieved."* (P31)

6 Discussion

This section seeks to interpret the results obtained from the proposed dynamics, initially seeking to answer the research question. The findings are subsequently discussed in relation to the related work.

6.1 Addressing the Research Question

To address the research question guiding this investigation, "How can LLMs support the phases of the Design Sprint?", the data collected indicate that ChatGPT, when guided by the proposed structured prompts, demonstrated the ability to support teams in the Understand, Diverge (Sketch), Converge (Decide), and Prototyping phases. This support contributed to the acceleration of artifact production and facilitated decision-making, highlighting the potential of AI to act as a facilitator within the Design Sprint context.

6.2 Comparison with Related Work

The results obtained are aligned with the findings of Torii et al. [18], which highlight the potential of generative AI to support creativity within the Design Sprint methodology. Similar to the present study, the mentioned research employed structured prompts to guide the process, although in a specific context of collaborative board game creation. This scenario illustrates the potential of utilizing AI in UCD and highlights the role of structured prompts as mediators of the creative process.

The work of Muehlhaus and Steimle [13] discusses the impact of prompt usage at each phase of UCD, emphasizing that structured and specific prompts are crucial to address the context of each step. This finding supports the approach adopted in the present study, which utilizes prompt engineering techniques throughout various activities. This approach may explain the participants' positive reception of the prompts' impact on artifact generation, as presented in the results section.

The strategy employed in this study, which includes post-prompt instructions aimed at reinforcing the need for human intervention after AI's response, aligns with the work of So [17], who advocates for integrating agile methodologies with machine learning without excluding active human participation. This approach encourages the combination of human critical judgment with AI-assisted support.

The positive outcomes observed in artifact generation during the study align with the results reported by Shaer et al. [16], who investigated the use of LLMs in collaborative ideation phases. Their analysis indicates that the presence of AI can enhance a team's creative capacity by offering contextualized suggestions, a characteristic also observed in the present research.

Finally, the open-ended responses indicated that, in certain situations, generative AI exhibited limited performance in image generation. This observation aligns with the findings of Kamnerddee et al. [9], who conducted a comparative study between a human-created prototype and an AI-generated one, revealing that the AI-produced prototype was simplistic and lacked elements that could engage users. These results underscore the importance of refining

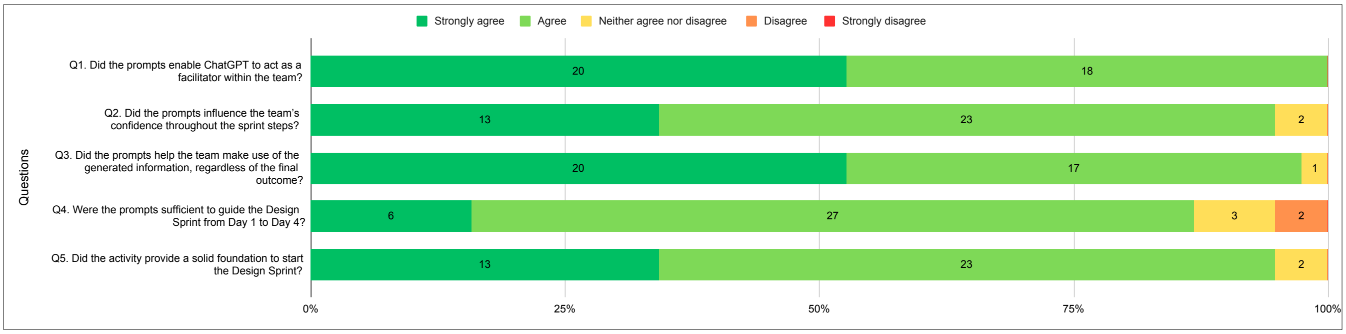

Figure 5: Individual Participant Feedback - Overall Evaluation of the Prompts.

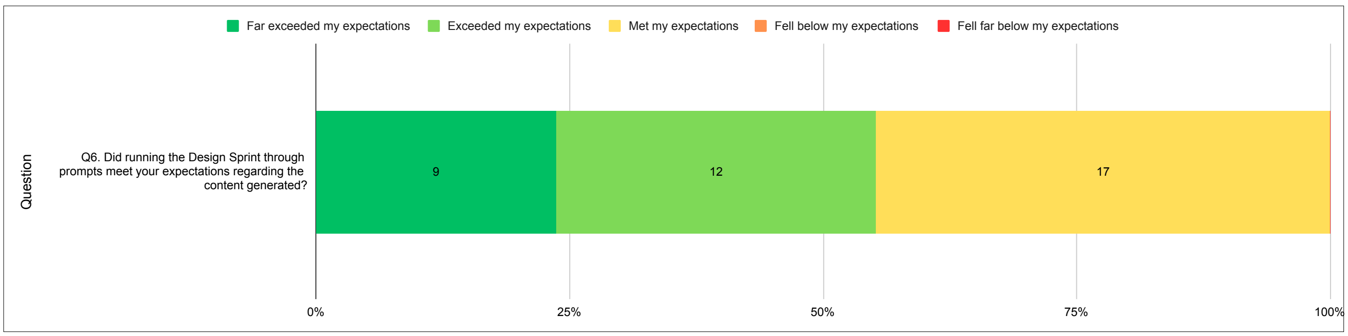

Figure 6: Individual Participant Feedback - Expectations Regarding the Generated Artifacts.

the prompts used for image generation throughout the process, as well as maintaining continuous human oversight of the generated artifacts.

7 Lessons Learned

This section aims to reflect on the overall experience of the study, highlighting its positive aspects, the challenges encountered, and the lessons that may guide future work.

- **Importance of Prompt Engineering:** The structuring of prompts proved to be fundamental for enabling ChatGPT to generate the desired artifacts. We observed that the quality of the responses is directly related to the content of the provided prompt.
- **Need for Prompt Improvement:** Based on the qualitative responses collected, it was identified that, in some cases, the prompts led to generic or contextually misaligned responses. For instance, the lack of specific information regarding the type of system (web or mobile) resulted in ambiguity. This scenario highlights the need for prompt revisions to increase their precision in future applications.
- **Human Supervision:** Although AI demonstrated the capability to automate Design Sprint activities, human intervention remains essential. The human critical characteristic is necessary to interpret the generated artifacts, validate content, and make decisions throughout the process.

8 Threats to Validity and Limitations

This section presents the threats to validity based on Wohlin [20], as well as limitations that should be considered when interpreting the results.

8.1 Internal Validity

- **Maturation:** Refers to the possibility that participants could be negatively affected by the intensive nature of the Design Sprint, leading to fatigue or demotivation throughout the activities, and consequently influencing the results.
Mitigation: To reduce this threat, the feasibility study was conducted over 14 days, allowing for the minimization of adverse effects arising from the intensity of the Design Sprint.
- **Instrumentation:** Refers to the possibility that the quality of the artifact prompts (the GenDS) could impact the results.
Mitigation: Although the artifact was created from scratch, the prompts were developed based on prompt engineering techniques to follow best practices in prompt construction. Additionally, the template of the updated 2024 version of the Design Sprint [26] was strictly observed, respecting the step-by-step instructions for creating each prompt.

8.2 Construct Validity

- **Confounding constructs and levels of constructs:** Refers to the possibility that participants had insufficient understanding of the Design Sprint or the GenDS. In this case,

the results could reflect difficulties in comprehending these elements rather than demonstrating the performance of the proposed approach.
Mitigation: To reduce this threat, prior training was provided on both the Design Sprint and the use of the artifact, providing participants with knowledge of the procedures before execution.

- **Evaluation Apprehension:** Involves the possibility of participant discomfort due to the evaluative nature of the study. **Mitigation:** To address this threat, the feasibility study was conducted remotely, without the direct presence of the researchers during execution. Moreover, participants were informed that personal identification data would be removed from the results and that they could withdraw from the study at any time without penalty.

8.3 External Validity

- **Interaction of selection and treatment:** The experiment was conducted in a controlled environment with computing students, which limits the generalization of the results. **Mitigation:** To mitigate this threat, 86,8% (33 of 38) of the participants had prior experience as industry professionals. Furthermore, a future study is planned to conduct the same investigation in an industrial context.

8.4 Limitations

- **Use of the free version of ChatGPT:** All participants utilized the free version of the model, which imposes limitations on access to advanced features. However, this scenario reflects the typical daily use of AI by academic students, making the experience representative of the context intended by the study.
- **Scope of automation up to prototyping:** The automation proposal was limited to the prototyping phase, so the testing phase was not included, as it would require user involvement, whereas the study's focus is to evaluate the feasibility of artifact generation by the team within the AI-assisted Design Sprint.
- **Absence of formal evaluation of generated artifacts:** A dedicated analysis of the quality of artifacts produced in each phase was not conducted, as the aim of the study is to investigate the feasibility of applying the Design Sprint with AI support, prioritizing participant feedback regarding interaction with the prompts.

9 Conclusion

The feasibility study enabled the evaluation of a preliminary set of prompts designed to automate the Design Sprint, facilitating a critical analysis of the impact of applying generative AI in the first four phases of the methodology. This analysis contributed to a deeper understanding of the benefits and limitations of this approach, providing valuable insights to support the continued research.

The results indicated a predominantly positive reception from the participants regarding the proposed dynamics. These findings suggest that, using the structured prompts developed, the ChatGPT model was able to support the teams in the Understand, Diverge (Sketch), Converge (Decide), and Prototyping phases. Besides, an acceleration in artifact generation and support for design decision-making was observed.

Given these findings, the continuation of the research is justified, with the allocation of resources for future investigations. As next steps, it is intended to refine the preliminary approach proposed and subsequently develop software that encapsulates this enhanced approach, incorporating the lessons learned.

ARTIFACT AVAILABILITY

The complete set of prompts, the presentations used during the trainings, and the artifacts resulting from the study are available at: https://doi.org/10.6084/m9.figshare.30293749

ACKNOWLEDGMENTS

We thank all the participants in the empirical study and the USES Research Group members for their support. We would like to thank the financial support granted by CNPq 314797/2023-8; CNPq 443934/2023-1; CNPq 445029/2024-2; CNPq - 313137/2025-0; CAPES - Financing Code 001; Amazonas State Research Support Foundation - FAPEAM - through POSGRAD 2025-2026. Funding was also provided by Project No. 017/2024 – DIVULGA CT&I/FAPEAM.

REFERENCES

[1] Chadia Abras, Diane Maloney-Krichmar, Jenny Preece, et al. 2004. User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications* 37, 4 (2004), 445–456.
[2] Liuqing Chen, Yaxuan Song, Jia Guo, Lingyun Sun, Peter Childs, and Yuan Yin. 2025. How generative AI supports human in conceptual design. *Design Science* 11 (2025), e9.
[3] Shiying Ding, Xinyi Chen, Yan Fang, Wenrui Liu, Yiwu Qiu, and Chunlei Chai. 2023. Designgpt: Multi-agent collaboration in design. In *2023 16th International Symposium on Computational Intelligence and Design (ISCID)*. IEEE, 204–208.
[4] Hubert Ekvall and Patrik Winnberg. 2023. Integrating chatgpt into the ux design process: Ideation and prototyping with llms.
[5] Angela Fan, Beliz Gokkaya, Mark Harman, Mitya Lyubarskiy, Shubho Sengupta, Shin Yoo, and Jie M Zhang. 2023. Large language models for software engineering: Survey and open problems. In *2023 IEEE/ACM International Conference on Software Engineering: Future of Software Engineering (ICSE-FoSE)*. IEEE, 31–53.
[6] Jeff Gothelf. 2013. *Lean UX: Applying lean principles to improve user experience.* " O'Reilly Media, Inc.".
[7] Andreas Holzinger. 2005. Usability engineering methods for software developers. *Commun. ACM* 48, 1 (2005), 71–74.
[8] International Organization for Standardization. 2019. *ISO 9241-210:2019 Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems.* Retrieved July 23, 2025 from https://www.iso.org/standard/77520.html
[9] Chayatad Kamnerddee, Pruet Putjorn, and Jittrapol Intarasirisawat. 2024. AI-driven design thinking: a comparative study of human-created and AI-generated UI prototypes for mobile applications. In *2024 8th International Conference on Information Technology (InCIT)*. IEEE, 237–242.
[10] Jake Knapp. 2024. *The 2024 Design Sprint by Jake Knapp.* Retrieved July 23, 2025 from https://miro.com/miroverse/the-2024-design-sprint/
[11] Jake Knapp, John Zeratsky, and Braden Kowitz. 2016. *Sprint: How to solve big problems and test new ideas in just five days.* Simon and Schuster.
[12] Yuwen Lu, Yuewen Yang, Qinyi Zhao, Chengzhi Zhang, and Toby Jia-Jun Li. 2024. AI assistance for UX: A literature review through human-centered AI. *arXiv preprint arXiv:2402.06089* (2024).
[13] Marie Muehlhaus and Jürgen Steimle. 2024. Interaction design with generative AI: An empirical study of emerging strategies across the four phases of design. *arXiv preprint arXiv:2411.02662* (2024).
[14] Hasso Plattner, Christoph Meinel, and Ulrich Weinberg. 2009. *Design thinking.* Springer.
[15] Sharif Uddin Ahmed Rana and Adrian David Cheok. 2025. Generative innovation: Leveraging the power of large language models for brainstorming. In *The Economics of Talent Management and Human Capital.* IGI Global, 175–192.

[16] Orit Shaer, Angelora Cooper, Osnat Mokryn, Andrew L Kun, and Hagit Ben Shoshan. 2024. AI-Augmented Brainwriting: Investigating the use of LLMs in group ideation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–17.

[17] Chaehan So. 2020. Human-in-the-loop design cycles–a process framework that integrates design sprints, agile processes, and machine learning with humans. In *International Conference on Human-Computer Interaction*. Springer, 136–145.

[18] Maya Grace Torii, Takahito Murakami, and Yoichi Ochiai. 2023. Lottery and sprint: Generate a board game with design sprint method on AutoGPT. In *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. 259–265.

[19] Ketmanto Wangsa, Ritesh Chugh, Shakir Karim, and Raj Sandu. 2022. A comparative study between design thinking, agile, and design sprint methodologies. *International Journal of Agile Systems and Management* 15, 2 (2022), 225–242.

[20] Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, Anders Wesslén, et al. 2012. *Experimentation in software engineering*. Vol. 236. Springer.

[21] John Zeratsky. 2016. *Sprint: Friday*. Retrieved July 23, 2025 from https://library.gv.com/sprint-week-friday-7f66b4194137

[22] John Zeratsky. 2016. *Sprint: Monday*. Retrieved July 23, 2025 from https://library.gv.com/sprint-week-monday-4bf0606b5c81

[23] John Zeratsky. 2016. *Sprint: Thursday*. Retrieved July 23, 2025 from https://library.gv.com/sprint-week-thursday-df8d7c8c0555

[24] John Zeratsky. 2016. *Sprint: Tuesday*. Retrieved July 23, 2025 from https://library.gv.com/sprint-week-tuesday-d22b30f905c3

[25] John Zeratsky. 2016. *Sprint: Wednesday*. Retrieved July 23, 2025 from https://library.gv.com/sprint-week-wednesday-900fe3f2c26e

[26] John Zeratsky. nd. *Design Sprint Template*. Retrieved Sep 26, 2025 from https://www.thesprintbook.com/template