From Prompts to Prototypes: Exploring ChatGPT in Collaborative Software Prototyping Education

Alexandre Pires
Federal Institute Farroupilha (IFFar)
São Borja, RS, Brazil
alexandre.2021307059@aluno.iffar.edu.br

Enzo Kleinpaul
Federal Institute Farroupilha (IFFar)
São Borja, RS, Brazil
enzo.2021310072@aluno.iffar.edu.br

Lorenzo Campos
Federal Institute Farroupilha (IFFar)
São Borja, RS, Brazil
lorenzo.68019@aluno.iffar.edu.br

Rafael Parizi
Federal Institute Farroupilha (IFFar)
São Borja, RS, Brazil
rafael.parizi@iffar.edu.br

ABSTRACT

Generative AI (GenAI) tools, such as ChatGPT, have the potential to accelerate software prototyping and enhance team collaboration by generating content, expanding ideation, and enabling faster feedback cycles. As prototyping is central to agile and user-centered design, helping teams validate ideas, refine requirements, and align solutions with user needs, GenAI may further amplify these benefits through iterative and collaborative support. However, how to effectively integrate GenAI into collaborative prototyping remains underexplored in Software Engineering (SE) education. This study explores how ChatGPT supports collaborative prototyping among student teams, focusing on interaction, organization, and decision-making. We conducted a case study with undergraduate teams using ChatGPT in project-based prototyping. Data were gathered through a questionnaire, prompt analysis, and observation. Team collaboration and prior experience influenced the use of ChatGPT. Jointly crafted prompts led to more relevant outputs, while simpler prompts resulted in reliance on generic responses. Students valued the tool for speeding up prototyping and ideation, despite challenges with prompt refinement and occasional misalignment. The study provides insights into the use of GenAI for prototyping in SE education, shedding light on its benefits and limitations for co-creation and team dynamics.

KEYWORDS

Software Prototyping, Generative AI, Team Collaboration, Software Engineering Education

1 Introduction

In Design and Product Discovery, prototyping helps validate ideas and ensure they meet user needs and strategic goals, directly contributing to software quality by aligning the final product with user expectations. User-centered methodologies and other active and collaborative learning practices, such as Design Thinking, support the creation of low-fidelity prototypes and interactive simulations that facilitate rapid iteration and early feedback collection [29], a recommended practice for improving software quality from the early stages of development. Prototyping also fosters communication among stakeholders, accelerates the tangible development of ideas, and guides decision-making based on user feedback [5].

Recent technological advances, particularly in Generative Artificial Intelligence, have created new opportunities to enhance this practice by supporting ideation, content generation, and interface simulation [28]. When applied to Software Engineering, GenAI tools can automate repetitive tasks, generate creative alternatives, and assist in refining early-stage solutions [21, 24, 31]. These capabilities make GenAI particularly promising for prototyping, as they enrich collaboration and promote shared understanding within development teams [27]. By accelerating prototype creation, GenAI enables more frequent feedback cycles, a cornerstone of software quality, as it supports the early validation of requirements and design decisions. However, despite growing interest, there is still limited research on how these tools can be meaningfully integrated into co-creation practices such as prototyping.

In this paper, we address the following research question: *"How does ChatGPT impact collaboration dynamics and software quality in student-led prototyping projects?"*. Our goal is to provide insights for both practitioners and educators, helping them understand whether GenAI tools enhance interaction, creativity, and shared decision-making, or introduce challenges such as reduced engagement and competitiveness, similar to those observed in gamification contexts[4]. We used ChatGPT[1] since it is one of the most used generative AI tools and is well-suited for educational use. We conducted a case study in which student teams integrated the tool into project-based learning activities to generate and refine prototypes. Our analysis examines how ChatGPT influenced team interaction, organization, and decision-making.

Our findings show that ChatGPT impacted collaboration in different ways, depending on students' technical backgrounds and team dynamics. While some teams actively refined and contextualized the AI-generated outputs through discussion, others relied more heavily on the content as produced. Prompt creation emerged as a central challenge that shaped both interaction and group cohesion. Reported benefits included (i) increased speed, (ii) improved organization, and (iii) expanded ideation, while challenges involved frequent prompt adjustments and occasional misalignment between the AI output and the teams' intentions [11].

The remainder of this paper is structured as follows: Section 2 presents related work on prototyping, requirements engineering, and generative AI in software development. Section 3 describes the methodology, including the course context, participants profile,

[1] chatgpt.com

data collection instruments, and evaluation procedures. Section 4 outlines our findings, while Section 5 discusses the results in light of related work. Finally, Section 6 presents our final remarks and directions for future research.

2 Background

2.1 Prototyping in Software Development

Prototyping is a very relevant task or activity in software development lifecycle since it helps to transforming abstract concepts into tangible artifacts that support exploration, validation, and decision-making [12, 30]. By engaging in prototyping, teams are able to test ideas, verify assumptions, and gather early feedback, helping ensure that products meet user needs and align with organizational goals [22, 29].

Beyond reducing uncertainty, prototyping fosters collaboration by providing a shared space for stakeholders to visualize and refine solutions [5]. However, aligning prototypes with user expectations often requires multiple iterations, which can make the process time-consuming and inefficient [10, 19]. Optimizing how teams conduct prototyping is therefore essential to accelerate development and reinforce user-centered practices.

2.2 Generative AI and Its Impact on Prototyping Dynamics

Recent advances in Artificial Intelligence, particularly in the subfield of Generative AI, have introduced new possibilities for software development [21, 24]. Trained on large-scale datasets, tools such as ChatGPT leverage deep learning to generate content such as text, images, and source code based on user prompts [15].

In prototyping contexts, teams can use natural language descriptions or simple prompts to generate wireframes, interaction flows, and even complex product simulations [25]. These prototypes are not intended to be fully functional, but serve as early-stage artifacts to test hypotheses with users and stakeholders before committing to full-scale development [31]. This approach enables early validation of ideas without requiring complete implementations [20].

By incorporating GenAI into the initial phases of product discovery, teams can more efficiently explore design and architectural alternatives and adapt features through agile feedback loops. In addition, integrating GenAI into prototyping may enhance collaboration and support shared knowledge construction [4]. However, it also poses risks: over-reliance on automation can reduce critical engagement [14], and poorly formulated prompts may generate low-quality outputs that require substantial refinement. Therefore, especially in educational settings, it is essential to assess whether GenAI supports co-creation or undermines active team participation and learning, while also ensuring that user expectations remain a central focus.

3 Research Methodology

We conducted a case study with undergraduate students from the Information Systems program at the Federal Institute Farroupilha (IFFar)[2]. The course is grounded in Design Thinking and Product Discovery principles and emphasizes collaborative and iterative practices for software ideation and prototyping, practices that are foundational to early, stage software quality assurance.

[2]www.iffar.edu.br

Our goal was to investigate how collaboration emerges and unfolds in educational contexts, within student teams, when ChatGPT is integrated into the prototyping process. To this end, we adopted a mixed-method approach, combining quantitative and qualitative data collection. We and encouraged participants to freely use ChatGPT to co-create and refine their software prototypes. We selected ChatGPT as the GenAI tool due to its capabilities in text generation, image creation, and coding assistance, which support rapid ideation, low-fidelity prototyping, and team-based decision-making in early-stage software design.

To guide our study, we formulated the following Main Research Question (MRQ): *"How does ChatGPT impact collaboration dynamics and software quality in student-led prototyping projects?"*. Based on this main question, we derived four supporting Research Sub-Questions (RQs), presented in Table 1.

This study followed the case study methodology proposed by Yin [33], which systematically investigates contemporary phenomena within real-life context. We structured the study into five main phases: (1) Planning, (2) Execution, (3) Data Collection, (4) Data Analysis, and (5) Results Presentation. Figure 1 illustrates the study's phases.

Phase 1 – Planning. The Planning phase involved defining the study protocol, preparing materials, and organizing the setting for the activity. We defined our case as a Software Engineering course offered within an undergraduate Information Systems program. The students enrolled in this course worked in teams previously formed during an earlier 18-week Design Thinking course, which ensured continuity in group dynamics. Our primary goal was not to create new interpersonal relationships or reorganize team composition, but to evaluate how teams that had already worked together integrated ChatGPT into their prototyping activities.

Before starting the study, we prepared a briefing outlining ChatGPT's potential applications for prototyping. We also developed two questionnaires, applied later in Phase 3 (Data Collection), which included both open- and closed-ended questions. Drawing inspiration from the Technology Acceptance Model (TAM)[9], we incorporated items to explore Perceived Usefulness (how students believed ChatGPT could enhance their team's prototyping performance and the quality of the resulting prototype) and Perceived Ease of Use (how intuitive and effortless they found the tool).

Table 1: Research Sub-Questions

#	Questions
RQ1	How do students perceive the usefulness, ease of use, and collaborative potential of ChatGPT, in educational context, during software prototyping activities?
RQ2	To what extent does the use of ChatGPT influence group decision-making during collaborative software prototyping activities?
RQ3	How do groups collaborate to build prompts in AI-supported software design?
RQ4	What do students perceive as the main benefits and challenges of using ChatGPT in software prototyping?

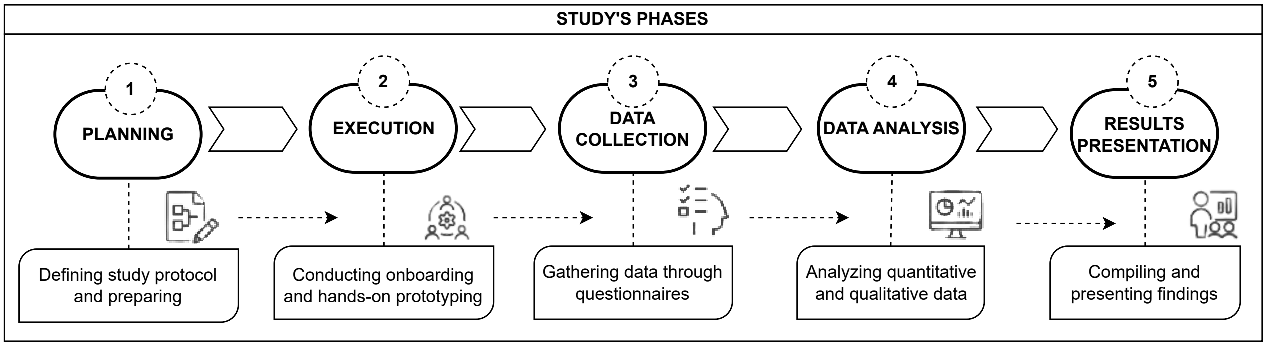

Figure 1: Study's Phases

Rather than applying TAM in its entirety, we adapted selected constructs to fit our educational and collaborative context, aiming to capture behaviours that might influence the teams' intention for future use. To ensure clarity and alignment with the study's objectives, we conducted a pilot test with a researcher in software engineering and human-computer interaction. Minor refinements were made based on their feedback.

Phase 2 – Execution. We conducted the Execution phase in two stages: (i) onboarding and (ii) hands-on prototyping with ChatGPT. In stage (i), we introduced the participants to the study objectives and to ChatGPT's role in supporting co-creation and prototyping. We gave a brief presentation to demonstrate the tool's capabilities and potential applications. In stage (ii), each team used ChatGPT to design and refine their software prototypes, iteratively improving the quality of their designs based on team discussions and AI-generated feedback. We encouraged the students to explore different prompting strategies and to observe how the tool could support interaction design, functionality ideation, and feedback refinement. Throughout the activity, we observed team behavior, communication dynamics, and the impact of AI integration on collaborative practices.

Phase 3 – Data Collection. During the Data Collection phase, the participants filled out the questionnaires. We also conducted direct observations of the groups during the activity, taking field notes to support the triangulation of qualitative data.

The first questionnaire included the Informed Consent Form (ICF) and a set of questions about the students' prior experience with ChatGPT and their perceptions of its use during the prototyping activity (see Table 2). The second questionnaire included ten items we designed to assess perceived ease of use, perceived usefulness, and the intention to reuse ChatGPT in similar future scenarios (see Table 3). We asked participants to rate the items on a 5-point Likert scale, ranging from 1 (Strongly disagree) to 5 (Strongly agree) [16]. In addition, we invited them to provide examples of the prompts they created and to describe how their teams collaborated in constructing them.

Phase 4 – Data Analysis. In the Data Analysis phase, we applied quantitative and qualitative methods. We analyzed the quantitative data from the questionnaire, using descriptive statistics (mean, median, standard deviation, minimum, and maximum). For the qualitative data, which included open-ended responses and examples of answers, a thematic content analysis was conducted based on the procedures described by Krippendorff [18]. We anonymized all collected data, including questionnaire responses and field notes, to ensure participant confidentiality. Then, we organized and cataloged the data in a digital repository to support analysis and ensure traceability.

We conducted the analysis using open coding. Two researchers independently examined the data to identify initial concepts and categories. To enhance the reliability of interpretations and mitigate potential bias, we held a cross-review session to compare codes, discuss emerging themes, and validate our findings. We resolved any disagreements through discussion until reaching a consensus on the final set of topics. This systematic approach allowed us to identify recurring patterns related to instructional strategies, group organization, and the benefits and challenges participants perceived when using generative AI tools during the prototyping activity.

Table 2: Individual Questionnaire: Questions for Evaluating the Use of ChatGPT During the Prototyping Activity

No.	Question	Answer Type
1	Had you used ChatGPT before for software prototyping purposes?	( ) Yes ( ) No
2	To what extent was ChatGPT responsible for the group's decisions?	( ) Low ( ) Medium ( ) High
3	How were the group's ideas combined with ChatGPT's suggestions?	( ) Directly generated by ChatGPT ( ) A mix of ChatGPT's suggestions and team members' own ideas
4	What were the main limitations encountered when using ChatGPT, as perceived by the group?	Open-ended
5	To what extent did the use of ChatGPT facilitate the completion of the task?	5-point Likert Scale (1 – Not at all to 5 – Very much)
6	What were the main aspects of ChatGPT that facilitated or hindered the process?	Open-ended
7	How did the group organize itself to create prompts for ChatGPT?	Open-ended

Table 3: Group Questionnaire: TAM-based questions to evaluate participants' perceptions of ChatGPT's usefulness and ease of use.

ID	Question
Q1	ChatGPT was easy to use for creating prompts and analyzing prototypes.
Q2	The use of ChatGPT significantly contributed to the development of the prototypes.
Q3	I would like to use ChatGPT again in future prototyping projects.
Q4	I plan to use ChatGPT in future collaborative projects.
Q5	ChatGPT's responses were reliable for supporting the prototyping activity.
Q6	ChatGPT facilitated collaboration among group members.
Q7	I identified barriers that hindered the use of ChatGPT during the activity.
Q8	Using ChatGPT reduced the difficulties encountered during prototype development.
Q9	The ideas suggested by ChatGPT significantly influenced the group's decisions.
Q10	ChatGPT helped generate better ideas than those initially proposed by the group.

These questions used a 5-point Likert scale (1 – Strongly disagree to 5 – Strongly agree) to evaluate participants' perceptions of ChatGPT's usefulness and ease of use.

Phase 5 – Results Presentation. In the Results Presentation phase, we compiled and organized the findings, which are presented in Section 4.

4 Results

This section presents the results of the case study conducted to investigate the use of ChatGPT in collaborative software prototyping activities, focusing on how the tool influenced collaboration, the design process, and the quality of the final prototype, which was generated in various different forms (html, image, documentation). We structured the section into five parts: Section 4.1 describes the participants' profiles, including their group allocation and previous experience with ChatGPT. Section 4.2 analyzes participants' perceptions of ChatGPT adoption using a TAM-inspired evaluation questionnaire, including its perceived usefulness in improving work quality. Section 4.3 explores the role of ChatGPT in influencing group decision-making processes and how these decisions impacted the overall prototype quality. Section 4.4 examines how ChatGPT impacted team organization and collaboration practices in pursuit of a higher-quality outcome. Finally, Section 4.5 highlights the main benefits (such as productivity gains and improved design quality) and challenges that participants experienced during the activity.

4.1 Participants and Previous Experience with ChatGPT

4.1.1 Participants' Profiling. The study involved 25 undergraduate students enrolled in a Software Engineering subject within an Information Systems program at IFFar. At the beginning, we organized the class into groups to collaboratively engage in a software prototyping activity. The course was built on a previous experience reported in Abich et al. (2024) [1], which examined the role of Design Thinking in fostering collaboration among SE students. That earlier study, conducted with 22 participants, focused on developing interpersonal and collaborative skills through iterative, user-centered design practices. While it effectively promoted teamwork and creative problem-solving, it lacked a dedicated prototyping phase, that is an essential step for bridging ideation and implementation in software development.

To address this gap, the current study introduced the use of GenAI (e.g, ChatGPT) to support the prototyping phase. Each group had full autonomy to define how they would interact with the tool, ranging from prompt formulation to the refinement and adaptation of AI-generated outputs. This flexibility led to diverse collaboration patterns and varying degrees of reliance on the AI across teams.

Table 4 summarizes the projects under development for each group and the number of participants involved.

4.1.2 Previous Experience with ChatGPT. We asked participants to report how frequently they used ChatGPT, using a 5-point Likert scale ranging from 1 ("Never") to 5 ("Very Frequently"). Figure 2 shows that most students reported regular to frequent use of the tool. Specifically, 36% selected "Frequently" and 28% chose "Very Frequently," indicating strong engagement. Additionally, 20% reported "Regular" use. In contrast, 12% indicated "Low Frequency," and only 4% had never used the tool. These results suggest that most participants were already familiar with ChatGPT before the activity, with higher usage observed among students with prior programming experience or involvement in similar projects.

When we asked participants about prior use of ChatGPT for software prototyping, 56% reported having used it for that purpose, while 44% had no prior experience. However, we did not identify any dependency between participants' technical backgrounds or professional areas and their responses.

4.2 Acceptance of ChatGPT for prototyping

To understand participants' perceptions of ChatGPT during the prototyping activity, we collected data through a questionnaire that included ten statements designed to assess perceived ease of use, perceived usefulness, and behavioral intention to use the tool (see Table 3). Table 5 presents the results, including descriptive statistics

Table 4: Summary of Group Interactions with ChatGPT

Group	Project	Participants
G1	Communication app for parents and teachers	6
G2	Interactive scheduling app for teachers	8
G3	Financial management system for family farming	5
G4	Awareness app for the correct use of pesticides	6

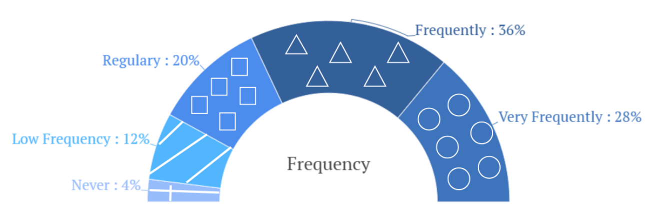

Figure 2: Frequency of use of ChatGPT by the participants.

(mean, median, standard deviation, minimum, and maximum) for each item, along with brief interpretive summaries.

The results indicate a positive perception of ChatGPT's ease of use (Q1) and a strong intention to reuse the tool in future prototyping projects (Q3). Participants also strongly agreed that ChatGPT supported collaboration within their teams (Q6). However, perceptions were more mixed regarding the reliability of responses (Q5), the quality of AI-generated ideas (Q10), and the tool's effectiveness in reducing task complexity (Q8) and maybe compromising the attendance of the user needs, as part of the software quality cycle.

4.3 The Role of ChatGPT in Group Decision-Making

This section examines how ChatGPT influenced decision-making processes within student teams during the prototyping activity. The analysis highlights varying levels of reliance on the tool and reveals participants' perceptions of its impact on the quality, structure, and direction of their work.

Table 6 summarizes the types of GenAI contributions identified at each level of influence. Figure 3 illustrates the three distinct levels observed across the groups. In low-influence scenarios (28%), students used ChatGPT primarily for early idea generation, but final decisions were made independently by the teams. In medium-influence cases (64%), ChatGPT supported the structuring of the prototyping plan and contributed to decision-making, with participants collaboratively adjusting and refining the AI's outputs. In high-influence scenarios (8%), ChatGPT had a more substantial impact on the teams' work, generating HTML code and UI mockups that were directly incorporated into the final prototypes.

Our analysis of the variation in influence levels reveals a direct impact on prototype quality attributes. In high-influence scenarios, where teams accepted AI-generated code with little to no modification, productivity increased, but concerns arose regarding the maintainability and reliability of the final artifact due to reduced critical evaluation. In contrast, medium-influence scenarios, characterized by collaborative refinement of the AI output, appeared to strike a more effective balance, enhancing the prototype's functional adequacy while preserving the team's control over its usability.

We also investigated how teams combined their ideas with AI-generated suggestions. As a result, 72% of participants reported blending ChatGPT's suggestions with their own ideas, while 28% accepted the AI-generated content directly, without major revisions. These patterns are consistent with the distribution of influence levels, suggesting that greater influence often correlates with less critical intervention.

To explore how participants perceived the impact of ChatGPT on their outcomes, we asked whether the results would have been the same without using the tool. Figure 4 shows that 64% of participants believed their results would have been different, while 36% stated that the outcome would likely have remained the same.

4.4 Group Organization

To understand how teams structured their interactions with ChatGPT, we analyzed how did the groups organize themselves to create the prompts. The results show that there was range of collaboration and delegation strategies, from centralized approaches, where one member typed while others made suggestions, to more participatory and iterative practices.

Across all four groups, a shared authorship model was common, in which one member acted as the typist while the rest of the team contributed ideas. This pattern was particularly evident in Groups 1 and 2. For instance, a student from G1 stated, –"*We gathered around one person; that person wrote while the others gave suggestions on what to write.*" Another participant from G1 described the process as a –"*brainstorming technique, analyzing all ideas and putting them in order (using AI).*"

In Group 2, the collective process of prototyping also fostered discussion on the complexity of crafting effective prompts. One student observed, –"*Everyone gave opinions, so the prompts were a bit confusing at first, but when we merged them with the idea we had previously elaborated, they became more functional.*" Another participant added, –"*After several failed attempts, we sat down and discussed the best way to ask [the prompt],*" highlighting a process marked by trial, refinement, and collective learning.

Group 3 adopted a hybrid strategy. Some prompts were created collaboratively, while others were developed individually and then integrated. A participant described this scenario as –"*two processes: one prompt created in group and another individually, then merging the ideas.*" Another participant stated that –"*only one member wrote, but the others helped with suggestions,*" indicating a division of roles with shared goal.

On the other hand, Group 4 demonstrated a slightly more structured and iterative approach. The students created prompts by using a combination of predefined design goals and group dialogue. One participant shared: –"*We decided together each line of the prompt to make it a clear command and get the best result possible.*" Other participants described the process as –"*assembling the ideas into a command so ChatGPT could generate code lines,*" and –"*one member typed the ideas that the group highlighted.*" Group G4 also mentioned

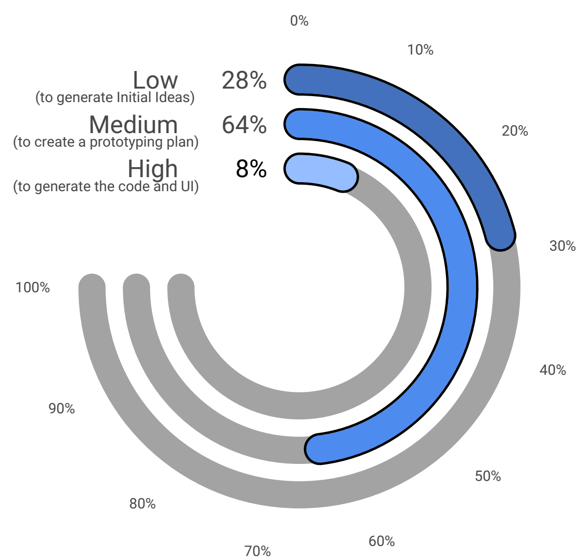

Figure 3: ChatGPT's Influence on Group Decisions

Table 5: Summary statistics and interpretation for each TAM-based question

Question	Mean	Median	Std. Deviation	Min	Max	Interpretation
ChatGPT was easy to use for creating prompts and analyzing prototypes (Q1)	4.00	4.0	0.82	3	5	High agreement among participants; Moderate variability in responses
ChatGPT significantly contributed to the development of the prototypes (Q2)	3.25	3.5	1.71	1	5	Moderate agreement among participants; High variability in responses
I would like to use ChatGPT again in future prototyping projects (Q3)	4.75	5.0	0.50	4	5	High agreement among participants; Low variability in responses
I plan to use ChatGPT in future collaborative projects (Q4)	3.25	3.5	2.06	1	5	Moderate agreement among participants; High variability in responses
ChatGPT's responses were reliable for supporting the prototyping activity (Q5)	2.75	3.0	1.50	1	4	Low agreement among participants; Moderate variability in responses
ChatGPT facilitated collaboration among group members (Q6)	4.50	4.5	0.58	4	5	High agreement among participants; Low variability in responses
I identified barriers that hindered the use of ChatGPT during the activity (Q7)	2.50	2.0	1.29	1	5	Low agreement among participants; Moderate variability in responses
Using ChatGPT reduced the difficulties encountered during prototype development (Q8)	3.25	3.0	1.50	2	5	Moderate agreement among participants; Moderate variability in responses
The ideas suggested by ChatGPT significantly influenced the group's decisions (Q9)	3.00	3.5	1.41	1	4	Moderate agreement among participants; Moderate variability in responses
ChatGPT helped generate better ideas than those initially proposed by the group (Q10)	2.50	2.5	1.29	1	4	Low agreement among participants; Moderate variability in responses

Table 6: Definition of degree of influence

#	Criteria
Low	Used just for support.
Medium	Used for support and changed some decisions
High	Completely changed the groups work directions.

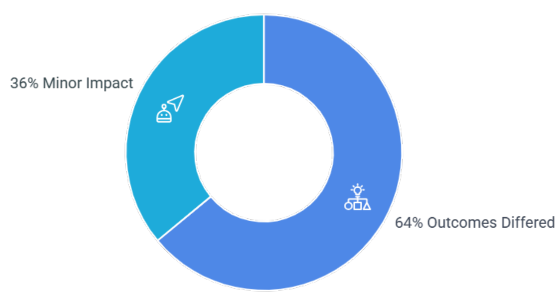

Figure 4: Participants' perception of ChatGPT's impact on final results

the use of Figma[3] and prior prototypes as a basis for guiding AI interaction.

Overall, the data suggest that the groups converged around three main organizational patterns:

A- **Centralized input with distributed ideation:** One member acted as the typist, while others provided real-time suggestions.

B- **Collaborative refinement:** Prompts were iteratively developed based on trial and error, group discussion, or the integration of individual drafts.

C- **Tool-supported co-creation:** Some groups used prior artifacts (e.g., Figma mockups) to inform and guide prompt formulation.

In addition to analyzing participants' perceptions, we examined the prompts created by each group to understand how different prompting strategies affected the quality of AI-generated outputs. Table 7 provides examples of the initial prompts formulated by the groups, along with a brief characterization of their prompting styles and interaction patterns.

An analysis of the prompts establishes a direct correlation between the chosen interaction strategy and the resultant software artifact's quality. Groups employing more sophisticated and iterative prompts (Groups 1 and 2) were effectively specifying quality requirements rather than simply issuing instructions. Through detailed descriptions of user flow or interface design, they directed the AI toward generating outputs with superior usability and functional adequacy. Conversely, the single-prompt strategies utilized by Groups 3 and 4, although faster, yielded generic outcomes that necessitated considerable manual post-processing to achieve an acceptable quality standard.

4.5 Perceived Benefits and Challenges of Using ChatGPT in Collaborative Prototyping

To better understand the perceived impact of ChatGPT during the prototyping activity, we analyzed participants' open-ended responses about the main benefits and challenges they encountered. The results show that while ChatGPT was widely valued for its performance, idea generation, and support in structuring the workflow, participants also identified notable limitations related to prompt specificity, limited contextual understanding, and the generic nature of some AI-generated content.

4.5.1 Benefits. The most frequently cited benefit across all groups was the speed and practicality of ChatGPT's responses. A participant from Group 1 noted, –"*The response time was very fast, and*

[3]figma.com

Table 7: Examples of Initial Prompts and Interaction Styles by Group

Group	Example of Prompt	Prompt Characteristics
G1	"*You are an expert in software prototyping and app development. I need you to create a prototype for an educational app that integrates communication between parents and teachers/tutors, including features like chat, activity posting, student profiles, and optional additional features for special needs students. Please include HTML/JS examples, system architecture, and API suggestions.*"	Very detailed and multi-faceted; combined requirements for design, implementation, and documentation; generation of prototypes focused on functionalities, generated in html; high complexity.
G2	"*Help me design a software prototype aimed at assisting teachers in organizing their schedules through an interactive calendar, with the help of a generative AI. Suggest interface features, data structures, and potential APIs.*"	Functional-focused initial prompt; evolved iteratively to include creative features like substitution management, room scheduling, and historical-based planning.
G3	"*Generate the construction of a software prototype related to financial management in family farming. Include main topics, problems, techniques used, collected data, and proposed solutions.*"	Direct and descriptive; aimed at system conception rather than detailed technical prototyping; minimal iterative refinement.
G4	"*Build a smartphone app prototype to teach and raise awareness among farmers about the correct use of pesticides, proper equipment usage, and correct waste disposal. Also suggest additional educational features.*"	Straightforward and goal-oriented; focused on educational content generation; later translated into a basic interface in HTML/JS/CSS.

it justified each answer." Similarly, a Group 4 member highlighted –"*the ease and speed with which it develops the software.*"

Another reported benefit was the tool's support in generating ideas that had not been initially considered. A participant from Group 3 observed that ChatGPT –"*delivered ideas we hadn't thought of,*" while another emphasized its usefulness in –"*organizing all data, structuring the work, and even setting the schedule.*"

Group 2 also emphasized ChatGPT's value in enhancing initial ideas, particularly in the early design stages: –"*It helped a lot to generate a more complete initial idea than the one we already had.*" The ability to produce varied suggestions was also welcomed, as one participant noted, –"*The diverse ideas were helpful.*"

In addition to speed and creativity, participants highlighted that the use of ChatGPT also contributed to improving the quality of the software. Organizing ideas, supporting task structuring, and defining the schedule helped the students make the process clearer and more efficient.

4.5.2 Challenges. Despite the reported benefits, participants identified several challenges. The most frequently mentioned issue was the need for highly precise prompts. As a student from Group 2 noted, –"*To get exactly what we wanted, everything had to be extremely detailed.*" Another participant added, –"*You need to spend time teaching it how to help you.*"

Some participants also pointed out a lack of contextual understanding or depth in the responses. A Group 4 participant remarked,–"*It gives you a direction to follow, but when it does things on its own, the content often lacks substance.*" Another noted that –"*it sometimes gives overly complex solutions for simple tasks.*"

Additional limitations included communication difficulties, –"*Understanding what we wanted was difficult,*" as one Group 2 student explained, and restrictions in the free version of ChatGPT: –"*The best version is paid, and it's hard to explain what you want before the session ends.*"

5 Discussions and Contributions

The use of ChatGPT for software prototyping has demonstrated potential for generating valuable insights in educational settings. In this study, we analyzed a set of related work to contextualize and support our findings [2, 3, 6–8, 13, 17, 26, 28, 32].

Table 8 summarizes each work's context, key findings, and its relation to our study.

Troya et al. [6] investigated the use of ChatGPT to support UML class diagram modeling in an academic setting. As in our study, students engaged in iterative prompt refinement and reflected on the challenges of using AI. While their focus was on identifying inconsistencies in AI-generated diagrams, we observed similar manifestations of critical thinking and collaborative decision-making during group interaction and prompt construction. In both contexts, guiding the AI response fostered shared understanding and collaboration. These findings suggest that, when embedded in structured learning activities, generative AI can promote not only task execution but also metacognitive development and peer learning.

Capozucca et al. [7] examined the use of ChatGPT to support formal specification writing with the B method. Their findings showed no overall improvement in students' performance, and in some cases, results worsened with AI use. However, students who critically evaluated the AI's output performed better. While their study focused on highly formal tasks, our investigation centered on collaborative prototyping.In this context, group interaction fostered critical reflection and contextual adaptation of AI suggestions. In both cases, guiding the AI and critically assessing its output were key to achieving meaningful results, and ensuring the quality of the final artifact.

Choudhuri et al. [8] conducted interviews to explore students' experiences using generative AI for learning and completing SE assignments. Among the topics discussed were motivations, usage contexts, and perceived benefits and limitations. One of their questions on the constraints of using GenAI closely aligns with our item on perceived challenges to ChatGPT use. In both studies, a recurring challenge was the difficulty of crafting effective prompts to obtain meaningful, context-aware responses. Our participants emphasized the need for highly detailed inputs, while Choudhuri et al. similarly found that students struggled to convey task context and requirements. Another shared concern was the lack of explanation in AI responses: both groups of students expressed frustration with the generic nature of outputs and the absence of clear reasoning behind the suggestions.

Table 8: Comparison between Related Work and Our Study

Reference	Context/Focus	Key Findings	Relation to Our Study
Troya et al. (2024) [6]	UML modeling with ChatGPT	Prompt refinement promoted reflection and collaboration	Similar iterative prompting and shared reflection
Capozucca et al. (2025) [7]	Formal specifications using B method	No improvement overall; critical stance improved results	Also found that critical evaluation leads to better outcomes
Choudhuri et al. (2024)[8]	Student perceptions of genAI use in SE	Prompt crafting and lack of explanation were key issues	Confirmed need for precise prompts and explanation
Al-Ahmad et al. (2024)[2]	ChatGPT in coding, bug fixing, documentation	GenAI saved time but needed manual validation	Showed how AI limitations fostered learning in groups
Waseem et al. (2023) [32]	Use of ChatGPT in semester-long project	AI bridged skill gaps but was used mostly for quick fixes	Reinforced AI's role in reflection, not just task support
Andersen-Kiel et al. (2024) [3]	Usability of ChatGPT in SE tasks (TAM-based)	Students found it easy but doubted reliability	Our students used ChatGPT more strategically in teams
Subramonyam et al. (2025) [27]	Professional use of genAI in prototyping	Prompt iteration and validation critical; bias risks observed	Parallels in prompt refinement and hallucination handling
Keuning et al. (2024) [17]	AI adoption in programming courses	Adoption depended on experience and course focus	Aligned with our findings on experience and task fit
Hanifi et al. (2023) [13]	Student use and hallucination reports	Hallucinations common despite high usage	Confirmed hallucinations and need for critical use
Ricardez et al. (2024) [23]	AI benefits in software dev and prototyping	AI supported creativity, speed, and organization	Matched reported benefits: speed, idea generation, structure

Al-Ahmad et al.[2] examined ChatGPT's impact on SE tasks such as coding, bug fixing, and documentation. While they emphasized gains in productivity and workload reduction, they also reported limitations in handling complex scenarios and providing context-aware solutions. In contrast, our study showed that in collaborative educational settings, these limitations became opportunities for learning. Rather than accepting AI outputs as final, students engaged in iterative refinement, discussed trade-offs, and co-developed prompting strategies, 72% of the teams adopted iterative prompting (Figure 3). This collective process, often absent in individual professional tasks, highlights ChatGPT's potential to foster not only technical assistance but also critical thinking and teamwork in educational environments.

Waseem et al. [32] examined ChatGPT's influence on university students during a three-month software development project, reporting gains in efficiency, collaboration, conceptual understanding, and interpersonal skills. While they highlighted ChatGPT's potential to bridge skill gaps, they also cautioned against overreliance. Although both studies demonstrate the tool's value in supporting development tasks, our findings reveal different usage patterns. Waseem et al. observed students using AI mainly to compensate for individual technical deficiencies, such as rapid code generation and reduced manual effort, whereas our participants adopted a more reflective and collaborative use. In group prototyping sessions, students refined AI outputs, discussed design trade-offs, and used the tool as a prompt for collective decision-making. Both studies acknowledged risks of dependency and code quality, but our results emphasize the potential of ChatGPT as a catalyst for critical thinking and team-based learning, which are essential for early-stage quality assurance.

Andersen-Kiel et al.[3] and Subramonyam et al.[28] highlighted how generative AI prompts critical reflection and the need for prompt refinement. While Andersen-Kiel et al. emphasized usability and students' doubts about reliability, our students demonstrated more strategic use of ChatGPT in collaborative prototyping sessions. Similarly, Subramonyam et al. discussed professional teams engaging in layered prompting and cross-validation. Likewise, it highlights how professional teams in design, engineering, and management conduct content-centered prototyping, exploring collaborative strategies for prompt writing, group decision-making criteria, and challenges related to model interpretability, points that resonate with our RQs on perceptions of utility and collaboration (RQ1), group decision-making (RQ2), and prompt co-construction strategies (RQ3); in our academic setting, student teams experienced similar challenges, including hallucinations and sensitivity to prompt phrasing, but leveraged these as learning opportunities through group-based iteration and refinement. These shared patterns reinforce the importance of prompt engineering as both a technical and pedagogical skill.

Keuning et al.[17] and Hanifi et al.[13] adopt a broader view of the software development cycle, focusing on risk analysis rather than collaboration. Their findings, while complementary, further support our results by highlighting the need for tools that enhance collaboration during early stages of software development. Students' acceptance of ChatGPT varied based on prior experience and course context, with concerns about learning erosion and technical inaccuracies being recurrent themes. In our study, students similarly recognized the risk of hallucinations. The students analyzed the generated response and considered the answer provided by the ChatGpt not good, requiring adjustments in the prompt, which generated new prototypes as shown in the Figure 5. However, the students emphasized the ChatGPT's usefulness for idea generation, rapid prototyping, and task organization. These results point to a convergence: while technical reliability remains a concern, ChatGPT's greatest value in educational prototyping lies in promoting creativity, structuring work, and fostering team reflection, particularly when embedded in well-designed collaborative activities.

While prompt engineering was a central element in students' interaction with ChatGPT, our analysis suggests that its use went

Figure 5: Prototype Generated by ChatGPT

beyond a mere technical activity. Teams engaged in prompt construction collaboratively, reflecting on the outputs and refining inputs iteratively, which fostered active dialogue and joint decision-making. This process stimulated the development of core competencies such as critical thinking, negotiation, and task organization, which are foundational practices for software quality assurance in early design stages. Rather than replacing human responsabilities, the use of GenAI acted as a fostering agent for co-creation and collective reasoning, especially in teams that treated the tool as a complement rather than a solution. These findings highlight that, under appropriate guidance, GenAI tools like ChatGPT can support not only prototyping efficiency but also meaningful learning in Software Engineering education.

Thus, our research contributes by analyzing these dynamics in a prototyping educational setting, where student engagement and a focus on collaboration emerge as central.

5.1 Takeaway Message for Educators

Based on our findings and the comparison with related studies, we highlight the following practical recommendations for educators seeking to integrate generative AI into prototyping tasks for software quality education:

- **AI as a Collaboration Catalyst:** ChatGPT can enhance team collaboration when used in structured group activities. Designing tasks that require shared prompt construction encourages communication and critical thinking.
- **Prompting is a Teachable Skill:** The quality of AI responses strongly depends on how prompts are written. Educators should explicitly teach prompt refinement techniques and provide examples of effective AI interaction strategies.
- **Scaffold the Use of AI:** Simply giving students access to AI is not enough. Clear guidelines, reflection questions, and formative checkpoints can help students use generative tools meaningfully rather than passively.
- **Encourage Critical Reflection:** Students should be prompted to assess the reliability and relevance of AI outputs. This reinforces metacognitive skills and helps them identify hallucinations and overgeneralizations.
- **Leverage AI for Ideation and Structuring:** ChatGPT was particularly useful for generating initial ideas, organizing work, and facilitating low-fidelity prototyping. Educators can design early-phase activities where these strengths are maximized.
- **Monitor Dependence and Guide Reflection:** While students appreciated AI assistance, some over-relied on default responses. Instructors should monitor this tendency and foster reflection to maintain autonomy and deepen learning.

6 Final Considerations

This study explored how integrating a Generative AI tool, specifically ChatGPT, can support collaboration in software prototyping activities within an undergraduate Software Engineering course. Using a case study methodology and mixed methods, we examined how student teams interacted with the tool to ideate, refine, and document software prototypes. The study was organized into five phases: planning, execution, data collection, data analysis, and results presentation. Participants were allowed to use ChatGPT freely throughout the activity. We collected data through an open-ended questionnaires, as well as prompt artifacts. This multi-source approach enabled us to capture not only perceptions of usefulness and ease of use, but also the collaborative dynamics that emerged during AI-supported prototyping.

Our findings indicate that ChatGPT was generally perceived as helpful for accelerating early design, but its effectiveness relied on the team's ability to craft precise prompts and collaborate openly. Teams that co-constructed prompts achieved more goal-aligned results, while challenges included repeated prompt refinement and occasional output misalignment. These insights underscore that effective use of GenAI tools in education depends not only on technical skill but also on collaborative engagement that treats AI as a co-creation partner.

This study provides practical insights for educators integrating generative AI into prototyping tasks. Our results suggest that tools like ChatGPT can enhance engagement, ideation, and collaboration, when supported by structured activities. Educators should design tasks that promote prompt refinement, shared decision-making, and critical reflection. Clear usage guidelines and formative assessments can help students treat AI not as a shortcut, but as a creative partner. Teaching prompting strategies explicitly can also strengthen both technical and collaborative skills in software design.

We addressed four types of validity in this study. Construct validity was reinforced through data triangulation (questionnaires, open responses, prompt artifacts) and a pilot with a senior researcher. Internal validity was supported by standardizing onboarding and ensuring consistent conditions. While external validity is limited to one course, the use of real teams and authentic tasks enhances applicability to similar settings. For conclusion validity, we applied established analysis methods and cross-checked interpretations. Future work could extend this investigation to other educational contexts, explore longitudinal effects on prompting and collaboration, compare AI model versions, and integrate AI into real-time prototyping platforms to support co-creation.

ARTIFACT AVAILABILITY

The full data package for this study is publicly available to ensure transparency and support replication. The artifacts, accessible at Zenodo (https://doi.org/10.5281/zenodo.15447395), include: (i) the anonymized raw data from all questionnaires; (ii) the data collection instruments (questionnaire templates); (iii) a complete and organized list of the prompts generated by each student group; and (iv) the scripts used for statistical analysis.

ACKNOWLEDGMENTS

We thank FAPERGS (Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul) for the financial support provided to Enzo Kleinpaul and Lorenzo Acosta. We also thank FAPERGS for the financial support to Rafael Parizi (grant nº 25/2551-0000925-1). We also acknowledge CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for the support granted to Alexandre Pires through the PIBIC 2024/2025 (Programa Institucional de Bolsas de Iniciação Científica) program.

REFERENCES

[1] Dariane Abich, Rafael Parizi, and Sabrina Marczak. 2024. Fostering Collaboration through Design Thinking: A Study among Software Engineering Students. In *Anais do XXXVIII Simpósio Brasileiro de Engenharia de Software* (Curitiba/PR). SBC, Porto Alegre, RS, Brasil, 532–542. doi:10.5753/sbes.2024.3544

[2] Ahmad Al-Ahmad, Hasan Kahtan, Luay Tahat, and Tarek Tahat. 2024. Enhancing Software Engineering with AI: Key Insights from ChatGPT. In *International Conference on Decision Aid Sciences and Applications (DASA)*. IEEE, 1–5.

[3] Noah Andersen-Kiel and Panagiotis Panos Linos. 2024. Using ChatGPT in Undergraduate Computer Science and Software Engineering Courses: A Students' Perspective. In *IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–9.

[4] Bilgram and Laarmann. 2023. Accelerating Innovation With Generative AI: AI-Augmented Digital Prototyping and Innovation Methods. In *In Proceedings of IEEE Conference*.

[5] Bogers and Horst. 2014. Collaborative Prototyping: Cross-Fertilization of Knowledge in Prototype-Driven Problem Solving. *In Proceedings of Journal of Product Innovation Management* (2014).

[6] Javier Cámara, Javier Troya, Julio Montes-Torres, and Francisco J Jaime. 2024. Generative AI in the Software Modeling Classroom: An Experience Report with ChatGPT and UML. *IEEE Software* (2024).

[7] Alfredo Capozucca, Daniil Yampolskyi, Alexander Goldberg, and Maximiliano Cristiá. 2025. Do AI assistants help students write formal specifications? A study with ChatGPT and the B-Method. *In Proceedings of arXiv preprint arXiv:2502.07789* (2025).

[8] Rudrajit Choudhuri, Ambareesh Ramakrishnan, Amreeta Chatterjee, Bianca Trinkenreich, Igor Steinmacher, Marco Gerosa, and Anita Sarma. 2024. Insights from the Frontline: GenAI Utilization Among Software Engineering Students. *In Proceedings of arXiv preprint arXiv:2412.15624* (2024).

[9] Fred D. Davis. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly* 13, 3 (Sep 1989), 319–340.

[10] Deshmukh and Wadhwa. 2007. A Meta Model for Iterative Development of Requirements Leveraging. In *In Proceedings of IEEE Conference*.

[11] Marco Gerosa, Bianca Trinkenreich, Igor Steinmacher, and Anita Sarma. 2024. Can AI serve as a substitute for human subjects in software engineering research? *In Proceedings of Automated Software Engineering* 31, 1 (2024), 13.

[12] Carolina Gill, Elizabeth Sanders, and Scott Shim. 2011. Prototypes as inquiry, visualization and communication. In *DS 69: Proceedings of E&PDE 2011, the 13th International Conference on Engineering and Product Design Education, London, UK, 08.-09.09. 2011*.

[13] Khadija Hanifi, Orcun Cetin, and Cemal Yilmaz. 2023. On chatgpt: Perspectives from software engineering students. In *2023 IEEE 23rd International Conference on Software Quality, Reliability, and Security (QRS)*. IEEE, 196–205.

[14] Xinyue Hao, Emrah Demir, and Daniel Eyers. 2024. Exploring collaborative decision-making: A quasi-experimental study of human and Generative AI interaction. *In Proceedings of Technology in Society* 78 (2024), 102662.

[15] Rishabh Jha. 2024. Incorporating Generative AI Into Human Resource Practices. *In Proceedings of SSRN Electronic Journal* (2024).

[16] Ankur Joshi, Saket Kale, Satish Chandel, and D Kumar Pal. 2015. Likert scale: Explored and explained. *British journal of applied science & technology* 7, 4 (2015), 396–403.

[17] Hieke Keuning, Isaac Alpizar-Chacon, Ioanna Lykourentzou, Lauren Beehler, Christian Köppe, Imke de Jong, and Sergey Sosnovsky. 2024. Students' Perceptions and Use of Generative AI Tools for Programming Across Different Computing Courses. In *Proceedings of the 24th Koli Calling International Conference on Computing Education Research*. 1–12.

[18] Klaus Krippendorff. 2009. *The content analysis reader*. sage.

[19] Kukreja and Boehm. 2013. Integrating collaborative requirements negotiation and prioritization processes: a match made in heaven. In *In Proceedings of ACM Conference*.

[20] Alicja Laskowska, Jesus Gorroñogoitia Cruz, Paweł Kędziora, Ilaria Lener, Bartosz Lewandowski, Cezary Mazurek, and Massimiliano Di Penta. 2014. Best practices for validating research software prototypes-MARKOS case study. In *eChallenges e-2014 Conference Proceedings*. IEEE, 1–9.

[21] et al. Mariana Coutinho. 2024. The Role of Generative AI in Software Development Productivity: A Pilot Case Study. *In Proceedings of arXiv preprint* (2024). arXiv:2406.00560 [cs.SE]

[22] Jürgen Münch, Stefan Trieflinger, and Bernd Heisler. 2020. Product Discovery – Building the Right Things: Insights from a Grey Literature Review. In *In Proceedings of IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*. 1–8.

[23] José Juan Ulín Ricárdez, Miguel Pérez Vasconcelos, Rosa Gómez Domínguez, and Fidelio Castillo Romero. 2024. Integration of Generative AI with ChatGPT in Software Development. *INNOVACIÓN Y DESARROLLO TECNOLÓGICO REVISTA DIGITAL* (2024).

[24] Patricia de Oliveira Santos, Allan Chamon Figueiredo, Pedro Nuno Moura, Bruna Diirr, Adriana C. F. Alvim, and Rodrigo Pereira Dos Santos. 2024. Impacts of the Usage of Generative Artificial Intelligence on Software Development Process. In *In Proceedings of the Brazilian Symposium on Information Systems* (Juiz de Fora, Brazil) *(SBSI '24)*. Association for Computing Machinery, New York, NY, USA, Article 65, 9 pages.

[25] Albrecht Schmidt. 2023. Speeding Up the Engineering of Interactive Systems with Generative AI. In *In Proceedings of the Symposium on Engineering Interactive Computing Systems* (Swansea, United Kingdom) *(EICS '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 7–8.

[26] Murillo Edson Souza and Murillo Edson Carvalho Souza. 2025. How Generative Artificial Intelligence Tools Can Improve the Quality of Software Produced by Software Development Teams. *In Proceedings of How Generative Artificial Intelligence Tools Can Improve the Quality of Software Produced by Software Development Teams* (2025).

[27] et al. Subramonyam. 2024. Content-Centric Prototyping of Generative AI Applications: Emerging Approaches and Challenges in Collaborative Software Teams. *In Proceedings of arXiv preprint* (2024). arXiv:2402.17721

[28] Hari Subramonyam, Divy Thakkar, Andrew Ku, Juergen Dieber, and Anoop K Sinha. 2025. Prototyping with Prompts: Emerging Approaches and Challenges in Generative AI Design for Collaborative Software Teams. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–22.

[29] et al. Szabluk. 2019. Experience design applied to research: an exploratory method of user centered research). *Revista UNISINOS* 8, 15 (2019).

[30] et al. Theis. 2022. A IMPORTÂNCIA DA PROTOTIPAGEM NO PROCESSO DE DESIGN E SUAS RELAÇÕES COMO MÍDIA DO CONHECIMENTO. *CIKI* 1, 1 (2022).

[31] et al. Ulfsnes. 2024. Transforming Software Development with Generative AI: Empirical Insights on Collaboration and Workflow. *In Proceedings of arXiv preprint* (2024). arXiv:2405.01543

[32] Muhammad Waseem, Teerath Das, Aakash Ahmad, Peng Liang, Mahdi Fehmideh, and Tommi Mikkonen. 2023. ChatGPT as a software development bot: a project-based study. *In Proceedings of arXiv preprint arXiv:2310.13648* (2023).

[33] Robert K Yin. 2015. *Estudo de Caso-: Planejamento e métodos*. Bookman editora.